

---

# Automatisk landformkartlegging med en kontekstuell relieffklassifisering

Bård Romstad



Hovedoppgave i geografi (geometrikk)  
Geografisk institutt  
Universitetet i Oslo  
Høsten 2001

---

## Forord

Hovedoppgaven er en del av cand. scient graden i geografi, studieretning geometrikk, ved Geografisk institutt, Universitet i Oslo. Oppgaven ble gitt av Dr. Bernd Etzelmüller som også har vært min veileder.

Oppgaven gjør bruk av feltdata som ble innsamlet på Svalbard sommeren 2000 i forbindelse med kurset GG 365, "Feltkurs i arktisk geomorfologi". Kurset var arrangert av Geografisk institutt og finansiert av Det matematisk-naturvitenskapelige fakultet, Universitetet i Oslo.

Jeg vil takke Dr. Bernd Etzelmüller for veiledning, støtte og entusiasme. Jeg vil også takke stipendiat Eva Heggem som jobbet sammen med meg på feltkurset og var med på både innsamlingen og prosesseringen av feltdataene. Øyvind Skaug takkes for tilrettelegging av programmet "massflux" og, sammen med Sigurd Haagenen, for interessante diskusjoner. Tilslutt rettes en takk til Dr. Jan Rasmus Sulebakk som tok seg tid til å kommentere det første utkastet til denne oppgaven, hans kommentarer var til stor hjelp.

## Sammendrag

Oppgaven omhandler klassifikasjon av terrengparametere, såkalt relieffklassifikasjon, for automatisk landformkartlegging. Som bakgrunn for dette ligger antagelsen om at det eksisterer en sammenheng mellom overflateprosess og overflateegenskaper som kan uttrykkes ved hjelp av topografiske parametere. Når rene attributtbaserte klassifikasjonsalgoritmer forsøkes brukt til strukturering av et datasett med relieffdata er resultatet ofte et sett av små spredte klasseforekomster i et påfallende kaotisk mønster. Siden geomorfologiske objekter, eller landformer, per definisjon er romlig koherente burde klassifikasjonsalgoritmen inkludere det romlige aspektet. Dette forsøkes gjort med en kontekstuell klassifikasjonsalgoritme som blir utprøvd på parametere derivert fra en 10×10 meter terrengmodell fra Brøggerhalvøya, Spitsbergen. Det vises at den kontekstuelle klassifikasjonen øker lesbarheten og anvendeligheten av den resulterende klassifikasjonen. Fordelingen av raskjegler forsøkes så forklart ved hjelp av forekomsten av visse terrengklasser. For å teste om dette forholdet kunne generaliseres til å bestemme fordelingen av slike landformer også i andre terreng ble den samme klassifikasjonen anvendt på en 30×30 meters terrengmodell fra Jotunheimen. Ved tolkning av flybilder ble det slått fast at også her sammenfaller de gitte landformene med forekomsten av de samme terrengklassene.

# Innhold

<b>Forord.....</b>	<b>1</b>
<b>Sammendrag .....</b>	<b>3</b>
<b>Innhold .....</b>	<b>5</b>
<b>1 Introduksjon.....</b>	<b>7</b>
1.1 Bakgrunn .....	7
1.2 Problemstilling og hypoteser.....	9
<b>2 Teori.....</b>	<b>11</b>
2.1 Geomorfometri .....	11
2.1.1 Spesifikk og generell geomorfometri .....	11
2.1.2 Topografiske mål.....	12
2.1.3 Sammensatte terrengattributter .....	13
2.2 Landskapsklassifisering.....	14
2.2.1 Landformavgrensing .....	14
2.2.2 Klassifikasjonsalgoritmer .....	15
2.2.3 Romlige hensyn.....	17
<b>3 Metoder .....</b>	<b>19</b>
3.1 Grunnlagsdata.....	19
3.1.1 Terrengmodeller .....	19
3.1.2 Felldata .....	20
3.2 Terrengparametere .....	22
3.3 Relieffklassifisering.....	23
3.3.1 Kontekstuell aggregering .....	24
3.3.2 Beregning av fasettparametere .....	26
3.3.3 Klassifisering .....	26
3.3.4 Kategorisering av landskapsklasser .....	27
<b>4 Resultater.....</b>	<b>29</b>
4.1 Brøggerhalvøya .....	29
4.2 Jotunheimen .....	34
<b>5 Diskusjon .....</b>	<b>37</b>
5.1 Sammenlikning av metodene .....	37
5.2 Romlig modellering av landformer og prosesser .....	38
5.3 Skala .....	39
5.4 Utnyttelse av høyere ordens parametere .....	40
5.5 Valg av terrengparametere .....	41
5.6 Generaliseringsgrad.....	42
5.7 Klassifikasjonsalgoritmer.....	42
<b>6 Konklusjon.....</b>	<b>43</b>
<b>Referanser .....</b>	<b>45</b>
<b>Vedlegg: Improving Relief Classification with Contextual Merging .....</b>	<b>49</b>

# 1 Introduksjon

## 1.1 Bakgrunn

I 1894 ga Albrecht Penck, Professor i geografi ved Universitetet i Wien, ut verket ”*Morphologie der Erdoberfläche*”. Her systemiseres tanker om klassifikasjon av landformer basert på deres kurvatur, eksposisjon, helning etc. i en hierarkisk struktur. Den kvantitative geomorfologien fikk et oppsving på midten av forrige århundre etter at Arthur Newell Strahler og hans studenter begynte med morfometriske studier basert på manuell analyse topografiske kart. Etter hvert øynet man også at slike analyser var velegnet til å automatiseres ved hjelp av den nye datateknologien som hadde begynt å vokse frem. I dag har IT revolusjonen for lengst nådd den kvantitative geografien. Geografiske informasjonssystemer (GIS) og økende tilgjengelighet til digitale terrengmodeller (DTM) har åpnet opp en ny verden av muligheter for kvantitativ landformsanalyse i den geomorfologiske forskningen (Pike, 1995; Pike, 2000). En av disse er muligheten for automatisk klassifikasjon av landformer (Weibel og DeLotto, 1988; Dikau, 1989; Sulebak et al., 1997). Når Penck beskrev jordoverflatens morfologi var nok målet å komme med en vitenskapelig, objektiv tilnærming til landskapsklassifikasjonen, ved å bruke kvantitative mål fremfor kvalitative og ofte kulturelt betingede beskrivelser. Lite ante han nok, at man i dag kan se for seg muligheten for å putte slike prinsipper inn i en datamaskin sammen med en terrengmodell og automatisk generere et geomorfologisk kart.

Automatisk klassifikasjon av terrengparametere for å identifisere og klassifisere ulike landformer, har et stort potensial innenfor geomorfologi. I geomorfologien er det gjerne sammenhengen mellom *form* og *prosess* som studeres og geomorfologiske prosesser er i stor grad styrt av topografien. Ny teknologi har gitt oss nye muligheter for å finne, analysere og kvantifisere slike sammenhenger mellom topografi og geomorfologiske prosesser. Det er dette som er hovedfokus innenfor fagretningen *geomorfometri* (Evans, 1972; Pike, 1995; Schmidt og Dikau, 1999). Geomorfometri, eller bare morfometri, defineres i denne oppgaven som *numerisk representasjon og analyse av topografien*. Det handler om å finne kvantitative mål på terrenget samt fornuftige metoder for analyse av disse. En grunnleggende antagelse er at det finnes en nær sammenheng mellom overflateprosesser og topografiske egenskaper, og at denne sammenhengen kan uttrykkes ved hjelp av topografiske parametere (Pike, 1988). Kombinasjoner av topografiske parametere definerer topografiske regioner etter en klassifikasjonsprosess. Det forventes at den romlige fordelingen av slike regioner er et uttrykk for hvilke geomorfologiske prosesser som har dominert eller dominerer i de ulike områdene.

Hvis en empirisk eller fysisk sammenheng mellom et sett topografiske parametere og en overflateprosess kan bli bestemt kan vi bruke denne sammenhengen til romlig modellering av geomorfologiske prosesser. Hvis vi videre antar at sammenhengen ikke er skalaavhengig er det mulig og opp- eller nedskalere en slik modell, i hvert fall innenfor en viss målestokk (cf. Etzelmüller et al., 2001).

Kommersielt tilgjengelig GIS programvare tilbyr i dag verktøy for beregning av topografiske parametere og klassifikasjon av multidimensjonale datasett. Det er likevel store utfordringer som må forseres før resultater fra slike analyser kan anvendes til geomorfologisk modellering av landformer. Disse utfordringene angår blant annet valg av modelltype og skala for representasjon av et kontinuerlig varierende terreng, valg og beregning av relevante parametere for en gitt applikasjon samt metoder for klassifikasjon, eller strukturering, av det kontinuerlig varierende terrenget til avgrensede geomorfologiske objekter (Mark, 1975; Weibel og DeLotto, 1988; Dikau, 1989; Moore et al., 1991; Schmidt og Dikau, 1999).

## 1.2 Problemstilling og hypoteser

I denne oppgaven er målet å se på ulike metoder for klassifikasjon av terrengparametere, og om resultatet av en slik klassifikasjon kan brukes til å identifisere områder der gitte geomorfologiske prosesser er dominerende.

Ved å anvende to ulike klassifikasjonsmetoder på en gridbasert terrengmodell fra Brøggerhalvøya, Spitsbergen, testes hypotesen:

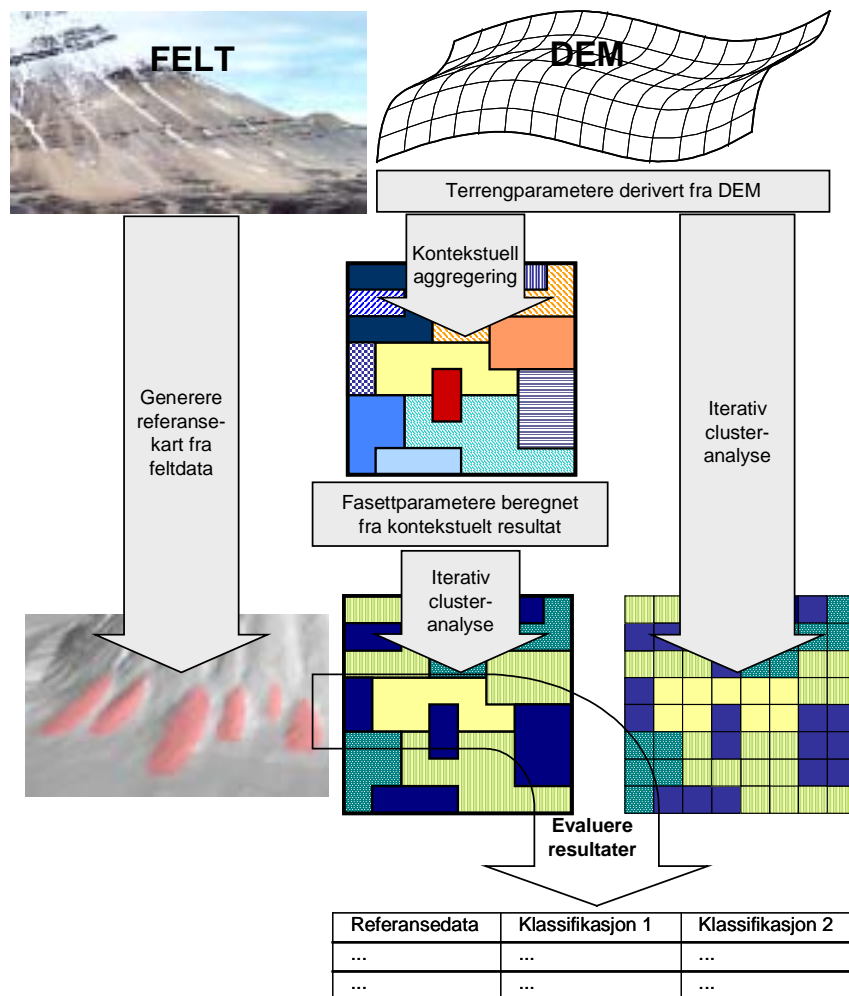
*”aktive skråningsformer kan identifiseres ved å klassifisere utvalgte terrengparametere og resultatet av en relieffklassifikasjon kan bli mer anvendelig dersom de opprinnelige punktene i datasettet først aggregeres opp til større romlige enheter”.*

I den ene klassifikasjonsmetoden gjennomføres først en generalisering av datasettet ved å aggregere de opprinnelige cellene til større romlige enheter. Deretter klassifiseres disse enhetene ved hjelp av en isocluster-algoritme. I den andre metoden klassifiseres de opprinnelige cellene direkte med den samme isocluster-algoritmen. Ved å sammenlikne resultatene med hverandre og med feltdata fra området kan de ulike metodene evalueres (se Figur 1).

For å se på den kontekstuelle klassifikasjonsmetodens forflyttbarhet er en annen hypotese utformet:

*”dersom klassifikasjonsmetoden er generelt anvendbar burde resultatene oppnådd i ett datasett kunne reproduseres i et annet datasett”.*

Denne hypotesen testes ved å anvende metoden på en terrengmodell fra Jotunheimen der både oppløsningen er lavere og relieffet er forskjellig fra Brøggerhalvøya. Dermed kan man se om eventuelle koblinger mellom klasseforekomster og geomorfologiske objekter funnet på Brøggerhalvøya også er tilstede her.



**Figur 1. Flytskjema for testing og sammenlikning av to ulike klassifikasjonsalgoritmer mot felldata.**



## 2 Teori

### 2.1 Geomorfometri

Geomorfologiske prosesser er hovedsakelig styrt av topografien. Økende tilgjengelighet til digitale terrengmodeller (DTM), og teknologi som er i stand til å prosessere disse (geografiske informasjonssystemer, GIS), har gitt nye muligheter for å definere kvantitative sammenhenger mellom topografi og prosesser (Beven og Kirkby, 1979; Beven et al., 1984; Moore et al., 1991; Moore et al., 1993a; Moore et al., 1993b). Anvendelsen av avansert teknologi for slike studier må basere seg på det teoretiske fundamentet i sin respektive disiplin, i dette tilfellet geomorfometri (Schmidt og Dikau, 1999). Geomorfometrien er i dag en syntese av matematikk, geovitenskaper, ingeniørfag og informatikk, men har røtter mer enn 150 år tilbake til Alexander Von Humboldt og senere tyske geografer (Pike, 2000). Disiplinen er stadig i endring ettersom teknologi og nye anvendelsesområder driver utviklingen fremover.

Det hersker en viss uenighet hvorvidt geomorfometrien kan kalles en egen vitenskap, og hvis ikke, hvilken vitenskap den i tilfelle tilhører, da den kombinerer elementer fra flere fagfelt (Pike, 1995). Det er likevel liten tvil om at geomorfometrien har mye å tilføre en rekke aktiviteter innen geomorfologisk forskning. Med den hurtige teknologiske utviklingen vi har hatt de siste tiår advarer Schmidt og Dikau (1999) mot at fokuset i den kvantitative geografien flyttes fra det vitenskapelige til det teknologiske. Evans, (1998) påpeker også at geomorfologisk forskning har endret seg fra å studere form for formens egen skyld til etter hvert å omhandle studiet av prosess for prosessens egen skyld. Kjernen i geomorfologien er interaksjonen mellom form og prosess og denne kjernen burde underbygges både ved analyse av den faktiske landoverflatens form og ved studiet av prosesser. Geomorfologien skal forsøke å forklare variasjoner i terrengoverflatens form og trenger derfor å identifisere viktige terrengparametere og etablere kunnskap om hvordan disse varierer (Evans, 1998).

#### 2.1.1 Spesifikk og generell geomorfometri

Evans (1972) skiller mellom *spesifikk* og *generell geomorfometri*. Den spesifikke geomorfometrien forholder seg til geometrisk analyse av spesifikke landformtyper. En søker å beskrive gitte landformer eller andre lokale fenomen i terrenget matematisk, men det foreslås ingen metode for en helhetlig og entydig landskapsklassifisering. Den generelle geomorfometrien derimot omhandler landformkarakteristikker anvendbare på enhver kontinuerlig terrengoverflate. Her etterstrebes en sammenfattende beskrivelse av overflaten

som ikke er begrenset til identifikasjon av gitte former. Selv om jeg i dette studiet fokuserer på identifikasjon av gitte former og bruker dette som et mål på klassifikasjonsmetodens egnethet så er det generelle geomorfometriske prinsipper som benyttes.

### 2.1.2 Topografiske mål

De fundamentale geomorfometriske konseptene er ifølge Mark (1975) de horisontale og vertikale skalaer av topografien. Dersom man ser for seg overflaten som en kombinasjon av periodiske funksjoner vil den korteste signifikante bølgelengden referere til de minste formene som kan identifiseres i en gitt terrengmodell. Dette kalles gjerne for terrengets *tekstur* (eng. *texture*). Den lengste signifikante bølgelengden refererer til terrengets hovedkomponenter, slik som fjell og daler, og kalles for *grain* (Evans, 1972; Mark, 1975). Amplituden forbundet med disse bølgelengdene beskriver terrengets relieff mens forholdet mellom små og store former, eller hvor mye av terrengets variasjon som forklares av de ulike bølgelengder, beskriver ruheten (Sulebak, 1999).

Alle topografiske mål tar utgangspunkt i høyden og hvordan denne varierer i rommet. I en terrengmodell kan høyden være representert som verdier i et regulært grid. Det medfører at høyden, som i virkeligheten varierer kontinuerlig over terrenget, diskretiseres til nivået av terrengmodellens oppløsning. Opplagt nok begrenses også skalaen det er hensiktsmessig å gjøre analyser på av denne oppløsningen. Mens høyden i seg selv er den grunnleggende topografiske parameteren er det muligens helningsvinkelen som er den viktigste. Dette fordi vinkelen kontrollerer gravitasjonskraften som er den drivende kraft for de fleste geomorfologiske prosesser (Evans, 1972). Helningsvinkelen er per definisjon den første deriverte av høyden og til forskjell fra bl.a. tekstur, grain, relieff og ruhet som beregnes uniformt for et avgrenset område, kan helningsvinkelen beregnes for hvert punkt på overflaten. Datastrukturer for representasjon av terrengoverflaten fører imidlertid stort sett til at helningen i praksis måles over en endelig avstand (Evans, 1972). Zevenbergen og Thorne (1987) beskriver hvordan bratteste gradient i hver celle kan beregnes for et regulært grid ved å tilpasse en matematisk flate til et 3×3 celler stort vindu rundt en sentercelle. Fra en slik flate kan også helningsretningen beregnes samt den andre deriverte av høyden, kurvaturen. Kurvaturen er et estimat av *endring i gradienten* eller *krumning* i hver celle. Den kan være konkav (positiv andre derivert) eller konveks (negativ andre derivert). Det finnes mange alternative måter å definere og beregne kurvatures komponenter på (se for eksempel Shary (1995)), men tre vel dokumenterte og utprøvde mål er *totalkurvatur*, *profilkurvatur* og

*plankurvatur*. Totalkurvaturen beskriver den generelle krumningen i et punkt. Selv om dette i seg selv kan være en interessant beskrivelse er det prosessmessig viktig hvilken retning krumningen forekommer. Profilkurvaturen beskriver krumningen langs helningsretningen og er en parameter som har vist seg å ha betydning for strømningsakselerasjon, erosjon/avsetningsrate, mens plankurvaturen beskriver helningen normalt på denne og har betydning for konvergent/divergent strøm og bakkefuktighet (Zevenbergen og Thorne, 1987; Moore et al., 1991). Selv om en positiv andre derivert gir en konkav form, hersker det en viss forvirring omkring hvorvidt et terreng med positiv kurvatur er konvekt eller konkavt. Innen geomorfologien er det imidlertid blitt en konvensjon å referere en positiv eller høy kurvatur til en konveks form siden man oftest betrakter endring langs en profil i skråningsretningen og konveksitet dermed impliserer en økning av gradienten (Evans, 1998).

### 2.1.3 Sammensatte terrengattributter

De topografiske parametrene beskrevet i forrige avsnitt er eksempler på det som gjerne kalles *primære attributter* (Moore et al., 1991). De primære attributtene kan kalkuleres direkte fra høydedata og har også blitt kalt *punktparametere* da de kan beregnes fra et punkts høydeverdi og punktets umiddelbare naboskap. Beveger vi oss over på et høyere romlig nivå kan vi beregne såkalte *sekundære* eller *sammensatte attributter*. Disse kan karakterisere den romlige variasjonen av en spesifikk prosess som forekommer i landskapet og er gjerne basert på en forenkling av de underliggende fysiske sammenhengene (Moore et al., 1991). Attributtene må beregnes i flere steg. Først må gjerne et *felt* eller *domene* defineres, deretter beregnes en verdi basert for eksempel på domenets størrelse, form eller topografi samt punktets plassering i domenet. Motivasjonen for bruken av slike attributter er altså å tilnærme en gitt fysisk prosess og spesielt innen hydrologisk modellering er dette mye anvendt. Beven og Kirkby (1979) beskriver en fysisk basert parameter kalt topografisk wetness index (WI). Denne er en funksjon av arealet som drenerer til et punkt (dreneringsarealet, eng. *flowaccumulation*) og helningsvinkelen i dette punktet. Indeksen er et mål på potensiell bakkefuktighet og kan beskrives med likningen

$$WI = \ln\left(\frac{A}{\tan(\beta)}\right) \qquad \textbf{Ligning} \qquad \textbf{1}$$

der  $A$  og  $\beta$  er henholdsvis dreneringsarealet og helningsvinkelen til punktet indeksen beregnes for. Økende dreneringsareal vil gi økt fuktighet, mens brattere helningsvinkel øker terrengets evne til å drenere og vil dermed føre til redusert fuktighet. Målet har vist seg å korrelere godt

med målte fuktighetsverdier i bakken (Beven og Kirkby, 1979; Beven et al., 1984; Moore et al., 1991; Moore et al., 1993a; Moore et al., 1993b). I en gridstruktur er det to ulike tilnærminger som kan brukes når man skal beregne dreneringsarealet. I den ene er målet å få algoritmen til å oppføre seg nokså likt vann som renner i strømmer på terrengoverflaten (bekker og elver). Når dreneringsveiene beregnes har en celle da kun lov til å drenere til nabocellen i retningen for størst helning (konvergent drenering). Den andre tilnærmingen er å la algoritmen oppføre seg mer som en tynn film av vann på eller i bakken og la en celle drenere til flere eller alle sine lavereliggende naboceller etter en fordelingsnøkkel (divergent drenering) (Budd og Warner, 1996). Moore et al. (1991) viser til at produktet mellom WI og plankurvaturen i noen studier har vist seg å korrelere bedre med fuktighetsmålinger i bakken en WI alene. Dette målet vil i tillegg til fuktighetstilgang og dreneringsevne ta hensyn til terrengets konvergerende eller divergerende effekt på strømminger. Det er likevel ting som tyder på at denne divergens/konvergens effekten ikke har like stor påvirkning på bakkens fuktighet ved ulike helninger, og at spesielt i flate områder domineres fuktighetsfordelingen av andre prosesser enn de som skyldes plankurvaturen (Sulebak et al., 2000).

## **2.2 Landskapsklassifikasjon**

Å klassifisere et landskap til funksjonelle morfologiske enheter som beskriver fysiske domener er noe som stadig gjøres innen geovitenskapene. Enten som et ledd i kartleggingsarbeid for kartleggingens egen skyld, men oftere til bruk i videre analyser. Konvensjonelt gjennomføres slikt arbeid gjennom feltundersøkelser eller ved tolkning av flybilder, noe som kan være svært tidsintensivt samt avhengig av en persons oppfattelse av et sett mer eller mindre kvalitative retningslinjer (Burrough et al., 2000). Den numeriske representasjonen av en overflates geometriske eller topografiske karakteristika, koblet sammen med forholdet mellom topografi, prosess og landform, gir oss muligheten til å definere et mer konsistent sett av retningslinjer for landskapsklassifikasjon.

### **2.2.1 Landformavgrensing**

Avgrensing av landformer på bakgrunn av deres geomorfometriske egenskaper blir av Pike (1988) gjort ved hjelp av det han kaller den *geometriske signatur*. Ideen er at det eksisterer et sett av parametere som beskriver den topografiske form godt nok til å kunne skille mellom ulike landskap. Weibel og DeLotto (1988) beskriver en automatisk landskapsklassifikasjon som en tredelt prosess. Først spesifiseres variable, deretter ekstraheres en geometrisk signatur og tilslutt klassifiseres denne. De tre stegene vil i praksis være svært avhengige av hverandre,

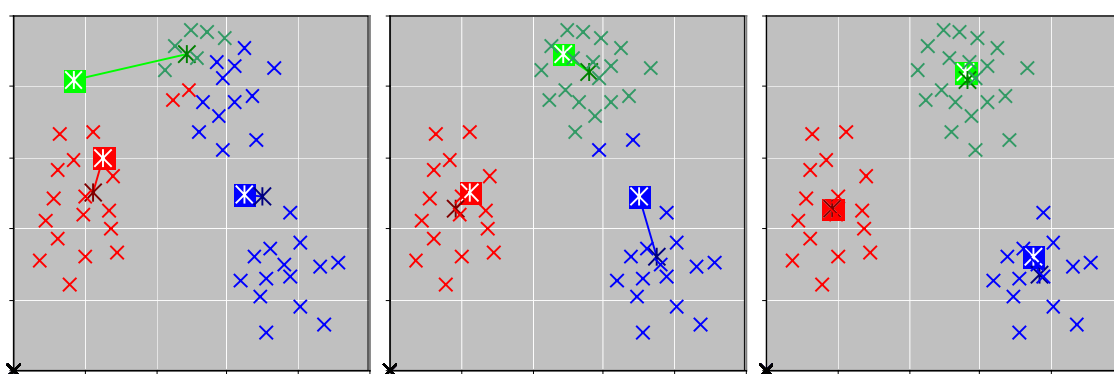
og til en viss grad også gli over i hverandre bl.a. fordi valg av variable vil påvirke hvordan signaturen kan ekstraheres og hva slags klassifikasjonsalgoritme som er hensiktsmessig. I tillegg må a priori kunnskap brukes i hvert steg for å sikre et tilfredsstillende resultat. Slik kunnskap er bl.a. kjennskap til hvordan terrengparametere beregnes og deres fysiske egenskaper. Dikau (1989) definerer et system for definisjon av geomorfografiske karakterer og deres romlige mønstre. Fra terrengmodellen formaliserer han definisjonen av objekter gjennom en hierarkisk oppdeling av romlige relieffenheter. Disse enhetene er (1) formfasetter som har homogen helning, retning og kurvatur, (2) formelementer med homogen plan og profilkurvatur og (3) relieff-former med homogen form. Relieffenhetene er definert kvantitativt som logiske kombinasjoner av DEM derivativer, og komplekse relieff-former kan genereres ved kombinasjoner av formelementer og fasetter.

### **2.2.2 Klassifikasjonsalgoritmer**

En klassifikasjon er en forenkling, eller generalisering, av attributter i et datasett. Hensikten kan være å gjøre informasjonen i datasettet lettere å oppfatte, å redusere en uhensiktsmessig stor datamengde eller for å definere klasser til bruk i videre analyse (Campbell, 1996). Det finnes mange metoder for klassifikasjon av data, men generelt søker man å strukturere datasettet til klasser med mindre forskjeller innen enn mellom hverandre. For et enkelt attributt kan klassifikasjonen være en ren oppdeling av dens verdier i intervaller, men med flere variable blir det nødvendig å finne en mer sammenfattende tilnærming. Fra fjernanalysen er bruken av statistiske metoder for klassifisering av multispektrale billed-data godt kjent. En vanlig inndeling av slike klassifikasjonsmetoder er i styrte og ikke-styrte algoritmer. I de styrte algoritmene defineres klassegrensene uavhengig av statistikken i datasettet ved at man manuelt definerer treningsområder der man vet at en gitt informasjonsklasse forekommer. De ikke-styrte algoritmene leter seg frem til naturlige grupperinger (clustere) i datasettet ved hjelp av en isodata (Iterative Self-Organising Data Analysis Technique) algoritme (Campbell, 1996). Begge typer av metoder har sine fordeler og ulemper. En styrt klassifikasjon kan kalibreres for en viss applikasjon og vil deretter kunne skape sammenliknbare klassifikasjonsresultater for ulike datasett, men man er avhengig av referansedata (eng. *ground true data*) for å definere klassegrensene. Med en ikke-styrt klassifikasjon kan man generere klasser som er optimalisert i forhold til datasettets fordeling i attributtrummet uten å vite noe om forholdene på bakken. Til gjengjeld vet man heller ikke hva de resulterende klasser representerer og sammenlikning mellom klassifikasjoner av ulike datasett er problematisk fordi klassegrensene vil endre seg fra klassifikasjon til klassifikasjon. De ikke-

styrte algoritmene er likevel populære ikke minst fordi de minimerer operatørens evne til å påvirke klassifikasjonsresultatet. Med generell kjennskap til typiske attributtkombinasjoner for ulike informasjonsklasser (informasjonsklassens signatur) samt en viss kjennskap til studieområdet vil man som oftest også kunne finne utvetydige koblinger mellom klassifikasjonsresultatet og informasjonsklassene.

Selve isodata algoritmen initieres gjerne ved at et gitt antall klassesentre spres tilfeldig ut i attributtrummet. For hver iterasjon beregnes den euklidske avstanden mellom hvert datapunkt og hvert klassesenter og datapunktene blir medlemmer av den klassen de er nærmest senteret til. Dermed kan nye klassesentre beregnes, basert på gjennomsnittet av alle klassemedlemmenes attributtverdier og prosedyren itereres til et konvergensminimum eller maksimalt antall iterasjoner blir nådd (se Figur 2). Man kan deretter tilegne datapunkter en klasse enten basert på den euklidske avstanden, som i iterasjonsprosessen, eller en annen form for avstand (Diagonal eller Mahalanobis) mellom pikselet og middelverdien for hvert cluster i det n-dimensjonale rommet (Burrough og McDonnell, 1998). Bruken av slike metoder for romlig strukturering av data har vært svært vellykket der dataene representerer naturlig skarpe (eng. *crisp*) klasser med liten overlap i attributt eller geografisk rom (Friedrich, 1996). Dette er ofte tilfelle for spektrale data fra luft- eller satellittbårne sensorer. Klassene som dannes fra slike data reflekterer ulike typer bakkedekke (eng. *landcover*) som er mer eller mindre distinkte. Overlapp mellom klassene er gjerne et skalaproblem (lav oppløsning fører ofte til blandingspikslar eller *mixles*) heller enn tvetydighet i klassedefinisjonene.



**Figur 2.** Eksempel på tre iterasjoner i en isodata analyse med tre klassesentre og to variable. De firkantede symbolene er nåværende klassegjennomsnitt mens pilen peker dit det nye klassegjennomsnittet vil være, basert på klassemedlemmenes fordeling.

### 2.2.3 Romlige hensyn

En iterativ clusteranalyse kan også anvendes på topografiske data, men slike data representerer stort sett et multivariabelt kontinuum som har få eller ingen distinkte klassegrenser. Dette er fordi det er en autentisk overlapp mellom ulike klasser i både rom og karakteristika (Friedrich, 1996). Overganger mellom ulike fasetter, eller landformer, i terrenget er som regel jevne. De ulike formene flyter over i hverandre og romlig avgrensing blir vanskelig. En enkelt landform består gjerne av flere elementer som hver for seg har ulike topografiske egenskaper. Noen av disse elementene kan godt være deler av flere landformer samtidig og hvorvidt et element med en gitt signatur er en del av en gitt landform er ofte avhengig av konteksten dette elementet befinner seg i. Dette faktum antyder at andre prosedyrer burde tas i bruk for å strukturere terrenget, metoder som tar hensyn til overlapp mellom klassene samt klassemedlemmenes kontekst (Weibel og DeLotto, 1988; Friedrich, 1996; Sulebak et al., 1997).

Irvin et al. (1997) og Burrough et al. (2000) beskriver bruken av kontinuerlige klassifikasjonsmetoder (eller "fuzzy set" metoder) for bedre å tilpasse klassifikasjonen til uskarpe (eng. *fuzzy*) klasser. Tanken bak slike klassifikasjonsmetoder er at klassemedlemskap ikke kan uttrykkes binært (ja/nei), men heller som en kontinuerlig medlemskapsverdi. Dermed kan en medlemskapsfunksjon mellom individuelle piksler og en klasse, heller enn det eksklusive medlemskapet, evalueres. Når denne funksjonen er beregnet kan pikslene defineres til en klasse (med en viss sannsynlighet). Man kan også definere mellomliggere, som befinner seg like nær, eller langt unna, to eller flere klasser. Irvin et al. (1997) sammenliknet en isodata og en kontinuerlig klassifikasjon av 6 parametere (høyde, helningsvinkel, helningsretning, profilkurvatur, plankurvatur og WI). De kom til at begge metodene produserte meningsfulle klasser, men at den kontinuerlige klassifikasjonen ga mer informasjon om klassenes karakter og variasjon enn isodata metoden. Burrough et al. (2000) bekrefter dette og konkluderer også med at metoden er robust også for større datasett og ulike relieff, men de påpeker at det er nødvendig å bruke romlig korrelerte inputdata da det er dette som i siste omgang sørger for at klassene er romlig konsistente. Kontinuerlige klassifikasjonsalgoritmer er forholdsvis enkle å inkludere i en clusteranalyse, men den flerdimensjonale informasjonen i resultatet er vanskelig å visualisere og en evaluering av klassene krever en omfattende forståelse for hvordan algoritmen virker og inputdataenes natur. Videre tar denne tilnærmingen i seg selv kun høyde for problemet med jevne overganger og overlapp mellom klassene i

egenskapsrommet og romlig korrelerte klasser er kun et biprodukt av terrengparametrene romlige egenskaper.

Friedrich (1996) beskriver en metode for aggregering av et multivariabelt datasett til homogene enheter. Grunnideen er at to *naboenheter* i terrenget som er svært eller helt like i egenskapsrommet kan bli slått sammen til en romlig enhet. På denne måten generaliseres datasettet så mye som mulig i det geografiske som, men beholder så mye som mulig av variasjonen i attributtrummet. Prosedyrene for bestemmelse av avstand mellom punkter i det multivariable rommet kan brukes som i clusteranalysen, men i tillegg må enheter som slås sammen altså være naboer i terrenget. Algoritmen minner om det som har blitt kalt "region growing" (Weibel og DeLotto, 1988) og virker som en slags resampling av data, men i stedet for å slå sammen celler som faller innenfor et rigid definert vindu brukes en prosedyre som for hver iterasjon slår sammen de to naboenhetene i datasettet som er likest hverandre. Resultatet er at enhetenes form og størrelse ikke lenger er definert av datastrukturen, men at de representerer mer eller mindre meningsfulle landskapsfasetter i terrenget. Det totale antall enheter og disses anvendbarhet i videre analyse vil derfor avhenge av størrelsen på studieområdet, hva slags tekstur og relieff det har, hva slags skala analysen foregår på samt variablene som brukes i aggregeringen. Friedrich konkluderer med at metoden virker lovende med hensyn til avgrensing av relieffenheter anvendbare innenfor for eksempel geomorfologi og jordfag.



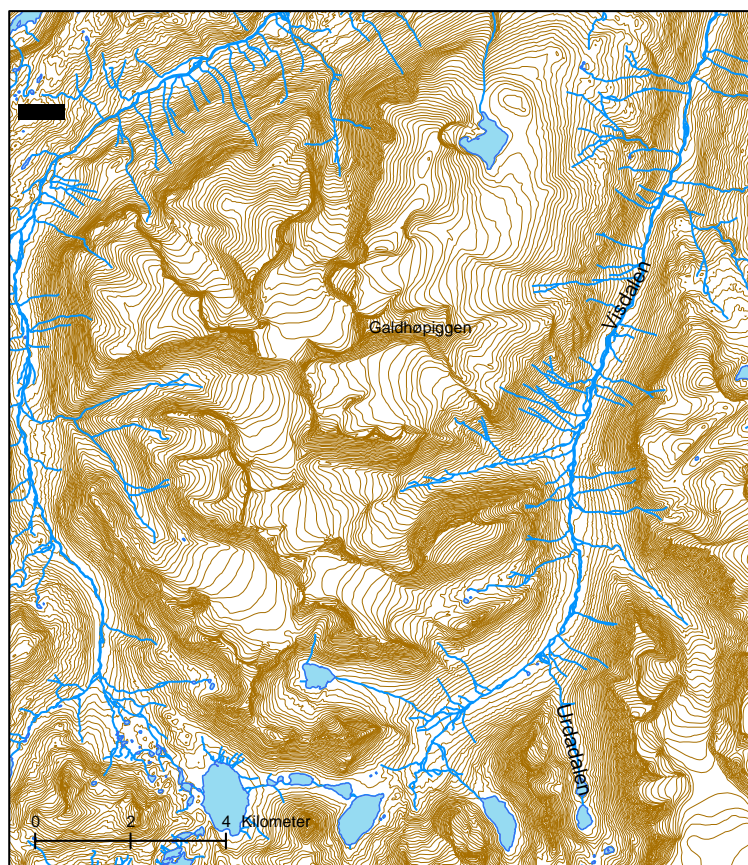
## 3 Metoder

Metodekapittelet består av tre deler. I den første delen beskrives alle grunnlagsdata, dernest følger en kort beskrivelse av terrengparametrene brukt i dette studiet og til slutt presenteres selve relieffklassifikasjonen.

### 3.1 Grunnlagsdata

#### 3.1.1 Terrengmodeller

Det ble brukt to ulike terrengmodeller i studiet, en fra Brøggerhalvøya, Svalbard, og en fra Jotunheimen. Den første modellen hadde en oppløsning på 10×10 meter, den var fremstilt av og tilgjengeliggjort fra Norsk Polarinstitut. Modellen fra Jotunheimen ble derimot fremstilt i forbindelse med dette hovedfagsarbeidet og beskrives derfor mer inngående. Modellen ble generert av forfatteren fra høydekoter med 20 meters ekvidistanse, elver, innsjøer og høydepunkter fra et kartblad i N50 kartserien til Statens kartverk (kartblad 15182, "Galdhøpiggen", se Figur 3). Dette ble gjort med en interpolasjonsalgoritme beskrevet av Hutchinson (1989) som er implementert i programvaren ArcInfo som funksjonen "topogrid" (ESRI, 2000). Algoritmen beregner først kurvaturer for hver celle basert på nærmeste høydekotes form, dette for å skaffe a priori kunnskap om hvordan høyden forventes å variere mellom kotene. Deretter interpoleres høydeverdier mellom kotene for hver enkelt celle. Man kan også informere algoritmen om dreneringslinjer (elver), innsjøer og høydepunkter. Den sørger da for at den interpolerte flaten drenerer konsistent, har flate innsjøer og skjærer høydepunktene dersom dette er mulig (konsistent drenering har prioritet). En viktig forutsetning for at en slik interpolasjon skal bli god er at alle koter er sammensluttet og at dreneringslinjer og innsjøer er entydig definert. Kartverket leverer ikke en slik kvalitet på sine digitale data og kartbladet måtte derfor manuelt gjennomgå og rettes opp for eventuelle feil. Den største jobben her er å lage et utvetydig elvenettverk. Man må sørge for at alle nodene i elvestykkene er koblet sammen og at elvestykkene peker i strømningsretningen. Algoritmen kan heller ikke ta hånd om elver som deler seg nedstrøms (braided streams) og forekomster av slike elveløp måtte derfor også fjernes manuelt. For å få så mye som mulig informasjon med ble også omkretsen til de innsjøene som hadde registrert høyde lagt til som høydekoter.



**Figur 3. Grunnlagsdata for beregning av DTM fra Jotunheimen.**

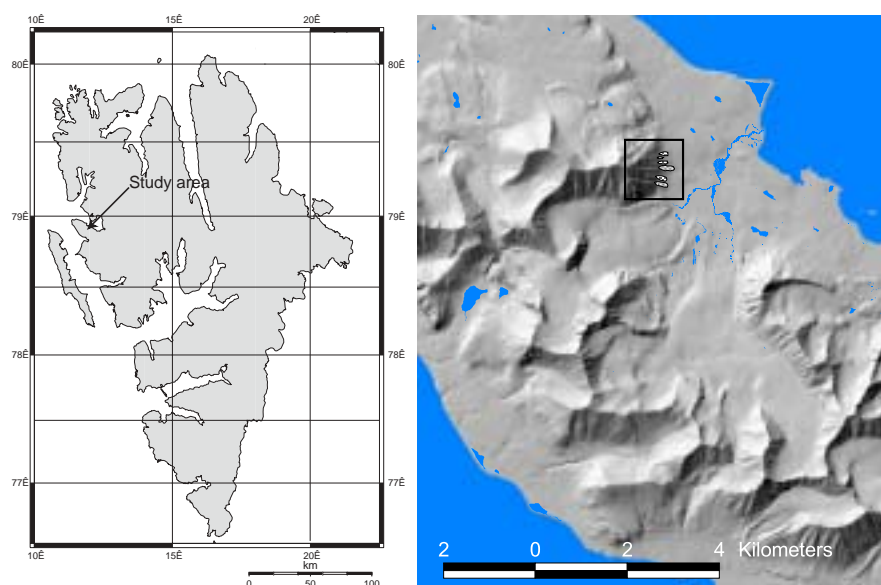
Det var ønsket å ha en oppløsning på 30×30 meter på terrengmodellen. I et terreng med så bratt relieff som det er i Jotunheimen viste det seg imidlertid at det ofte var mindre enn 30 meter mellom høydekontene. Algoritmen produserte derfor et bedre resultat (basert på terrengmodellens avvik fra kjente høyder) når celledørrelsen ble redusert til 10×10 meter for så å bli resamlet opp til 30×30 meters celledørrelse med en bikubisk resampling.

### 3.1.2 Felldata

For å kunne evaluere kvaliteten av et klassifikasjonsresultat er det essensielt å ha tilgang på referansedata som reflekterer de "sanne" forholdene på bakken. Etter en klassifikasjon kan resultatet sammenliknes med disse referansedataene og man kan få et mål på klassifikasjonens egnethet.

Felldata ble samlet inn som en del av kurset GG 365 "Feltkurs i arktisk geomorfologi" arrangert av Geografisk institutt, UiO, i august 2000, under ledelse av Professor J.L. Sollid og Dr. B. Etzelmüller. Mesteparten av feltkurset foregikk i nærheten av Ny-Ålesund på

Brøggerhalvøya, Spitsbergen (se Figur 4). I løpet av feltkurset ble det gjort oppmålinger av fire taluskjegler langs Scheteligfjellets østskråning. Taluskjegler, eller raskjegler, er koniske ansamlinger av løsmateriale som har forvitret fra skråningen over kjeglen. Målingene ble foretatt ved hjelp av målbånd, kompass og altimeter. Kjeglenes lengde, bredde, retning og høydeforskjell ble registrert og det ble tatt fotografier av området. Ved hjelp av flybilder, et digitalt grunnlagskart (høydekoter og elver) samt feltobservasjonene kunne formene plasseres på kartet. Talusenes toppunkt kan defineres som der terrenget endrer seg fra en konkav, til en konveks plankurvatur. Dette punktet var lett gjenkjennelig fra høydekotene i grunnlagskartet. Deretter kunne resten av kjeglen plasseres ved hjelp av retnings-, lengde-, bredde- og høydemålingene. Etter å ha plassert de fire oppmålte talusene på Scheteligfjellet ble det resulterende kartet sammenliknet med flybilder og på bakgrunn av dette kunne ytterligere to taluser i den samme skråningen kartlegges.



**Figur 4.** Feltområdet på Brøggerhalvøya, Spitsbergen. Scheteligfjellets østskråning og de registrerte landformene er innfelt på figuren til høyre.

Kartet ble tilslutt lagret på gridformat med celledørrelse tilsvarende terrengmodellen fra samme område (10×10 m). Vel så viktig som de kartlagte formene var også kunnskapen om at det i resten av det undersøkte området ikke fantes flere taluskjegler enn de som var kartlagt, derfor ble kartets totale utstrekning begrenset til dette området (se Figur 5).



**Figur 5. Bilde av feltområdet og terrengmodell med de registrerte taluskjeglene.**

### **3.2 Terrengparametere**

Seks topografiske parametere ble beregnet fra terrengmodellen. Disse var helningsvinkel, total-, profil- og plankurvatur, topografisk wetness index (WI) og tilslutt produktet av WI og plankurvaturen (WIplan).

Helningsvinkelen ble i dette studiet beregnet fra et naboskap på  $3 \times 3$  celler rundt den fokale cellen. Algoritmen er beskrevet i Burrough (1986) og implementert i GIS programvaren ArcInfo som funksjonen "slope" (ESRI, 2000). Resultatet gis i grader.

Total-, profil- og plankurvatur ble også beregnet fra et naboskap på  $3 \times 3$  celler. Algoritmen er beskrevet både av Zevenbergen og Thorne (1987) og av Moore et al. (1991) og er implementert som "curvature" i ArcInfo (ESRI, 2000). Resultatet gis i  $1/100$  z-enheter og positive verdier representerer en konveks form.

Topografisk wetness index (WI) er som kjent en funksjon av dreneringsarealet til og helningen i et punkt og ble beregnet som i Ligning 1. For beregning av dreneringsarealet ble det til dette studiet funnet mest hensiktsmessig å bruke en algoritme med divergent drenering. Dette fordi en slik algoritme er mindre sensitiv til feil i terrengmodellen og da selve hydrologien egentlig var underordnet hva indeksen kunne si om en celledes topografiske karaktertrekk. Til beregning av dreneringsarealet ble det brukt et spesiallaget dataprogram (Bjørnstad et al., 1998) som fordeler dreneringen fra en celle til alle laveliggende naboceller, j, som i Ligning 2

$$Andel_j = \frac{\frac{dZ_j}{D_j}}{\sum_{i=1}^n \frac{dZ_i}{D_i}}$$

Ligning

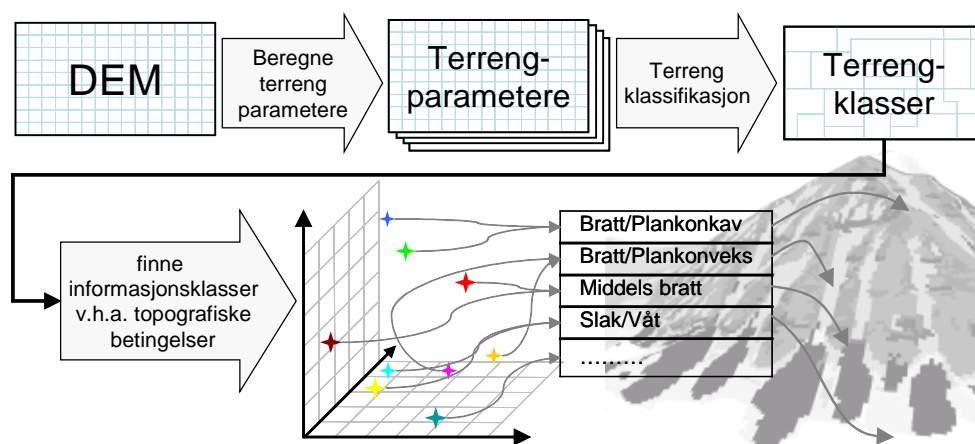
2

der  $dZ$  er høydeforskjellen mellom givercelle og mottakercelle,  $D$  er avstanden mellom disse og  $n$  er antall lavereliggende celler.

Produktet mellom WI og plankurvatur (WIplan) ble beregnet ved å multiplisere WI med den negative plankurvaturen. Parameteren forventes da å korrelere positivt med kombinasjonen av terrengets konvergens og potensielle tilgang på fuktighet. Høye positive verdier antyder altså både høy WI og konvergerende kurvatur. De laveste verdiene forekommer imidlertid også der WI er høy, men kurvaturen er konveks.

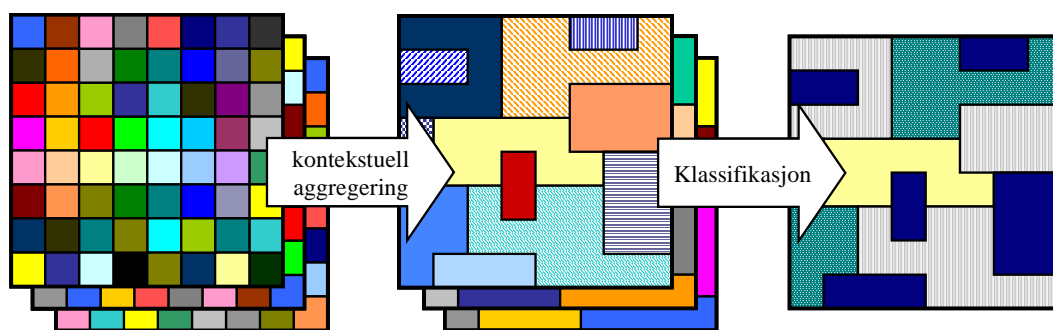
### 3.3 Relieffklassifikasjon

Metoden er skissert i flytskjemaet på Figur 6. Topografiske parametere beregnes fra terrengmodellen og brukes som input i en klassifikasjonsprosess før klassene tilegnes til landskapstyper ved å evaluere hver klasses plassering i attributtrommet. Selve terrengklassifikasjonen er en tredelt prosedyre skissert i Figur 7. Først benyttes den kontekstuelle aggregeringsalgoritmen beskrevet av Friedrich (1996) som aggregerer landskapet til større enheter. For hver av de resulterende landskapsfasettene blir nye terrengparametere beregnet. Disse parametrene er grunnlagsdata for en iterativ clusteranalyse slik at fasettene kan klassifiseres til mer generelle landformtyper



**Figur 6.** Flytskjema for metoden i skråningskartleggingen. Terrengparametere beregnes fra høydemodellen, disse klassifiseres og de resulterende klassene tilegnes til landskapstyper ved å evaluere deres plassering i attributtrommet.





**Figur 7. Selve terrengklassifikasjonen er tredelt: De originale parametrene generaliseres med en kontekstuell aggregering. For hver av de resulterende enhetene beregnes nye parametre som så klassifiseres til mer generelle klasser.**

Terrengmodellene og beregning av relieffparametere fra disse er allerede berørt i avsnitt 3.1 og 3.2 så nedenfor følger en nærmere beskrivelse av de ulike stegene i selve terrengklassifikasjonen og klassekategoriseringen.

### 3.3.1 Kontekstuell aggregering

Aggregereingsalgoritmen beskrevet av Friedrich (1996), ble anvendt på datasett der helningsvinkel, profilkurvatur, plankurvatur og WI var beregnet som attributter for hver celle. Av de seks terrengparametrene som var beregnet fra terrengmodellen var disse fire de minst korrelerte og ble derfor ansett som best egnet til å diskriminere ulike terrengtyper. Før dataene ble puttet inn i aggregeringsalgoritmen ble de normalisert og vektet. Kurvaturene ble normalisert mellom -2 og 2, helningsvinkelen mellom 0° og 45° og WI mellom 0 og 12. For denne applikasjonen ble helningsvinkelen ansett som en viktigere avgrensingsparameter enn kurvaturmålene og ble derfor vektet med 2, mens WI ble ansett som viktigere enn helningen igjen og ble vektet med 4. Kurvaturmålene, som også er mer sensitive til feil og artefakter i terrengmodellen (Evans, 1998), behold altså sin opprinnelige vekt på 1.

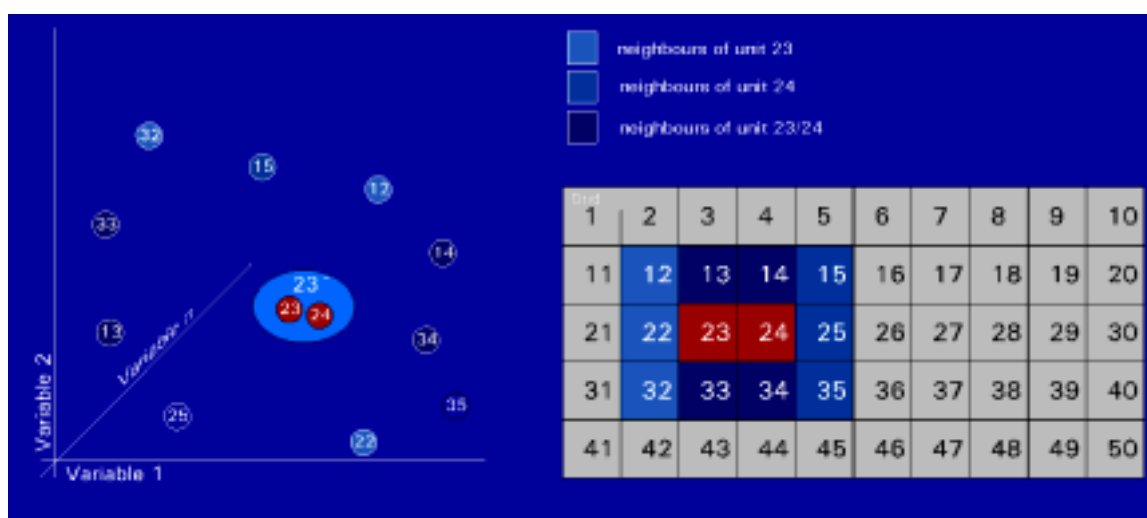
Algoritmen er basert på avstandsvektoren i attributtrummet mellom celler som er naboer i terrenget. Til å begynne med er hver celle oppfattet som en unik enhet, for hver av enhetene beregnes avstanden til alle naboer som i Ligning 3

$$\vec{v} = \sum_i^n \sqrt{(a_i - b_i)^2}$$

**Ligning 3**

der  $\vec{v}$  er vektoren mellom enhet  $a$  og  $b$ ,  $n$  er det totale antall variable og  $a_i$  og  $b_i$  er verdien av variabel  $i$  i enhetene  $a$  og  $b$  respektivt. Vektorene lagres i en matrise etter hvert som de

beregnes og når avstanden mellom alle naboer i datasettet er kjent slås de to naboene sammen med den minste vektoren sammen. Disse utgjør nå en ny enhet, med ny middelværdi og nye naboer (se Figur 8). Avstanden mellom den nye enheten og de nye naboene må beregnes på nytt, avstandsmatrisen oppdateres og prosedyren gjentas til et avbruddskriterium er nådd. Dette avbruddskriteriet vil være bestemt av hvilken grad av generalisering man ønsker, eller med andre ord, hvor mange enheter man vil ende opp med. En passende generaliseringsgrad vil være avhengig av datasettets oppløsning i forhold til skalaen til objektene man ønsker å analysere, jo høyere oppløsning på grunnlagsdataene jo større generalisering kan man tillate. Terrengets tekstur og relieff påvirker også hvor stor generalisering som er ønskelig. Dette fordi landformer i et jevnt terreng per definisjon er mer homogene enn de er i et røffere landskap. Dette fører til problemer dersom det finnes landskap med ulike relieff innen studieområdet. Algoritmen vil da ha en tendens til å aggregere opp ulike landformelementer i de jevne områdene *før* celler innen ett og samme element slås sammen i det røffere landskapet.



**Figur 8. Den kontekstuelle aggregeringsalgoritmen. De to nærmeste enhetene slås sammen (23 og 24) og nye avstander må beregnes mellom den nye enheten (23') og dennes naboenheter (12–15, 22, 25 og 32–35) (fra Friedrich, 1996).**

Selve aggregeringsalgoritmen er implementert i et program kalt IVHG beskrevet av Friedrich (1996). I det originale programmet bestemte man på forhånd det totale antall enheter man ønsket å ende opp med og dette ble brukt som avbruddskriterium. Siden man kjenner antallet celler i datasettet kan dette også beskrives som en prosentvis grad av generalisering i det geografiske rom. For å gjøre avbruddskriteriet mindre sensitivt for terreng med ulikt relieff

var det derimot ønskelig å bruke graden av generalisering i attributtrummet som avbruddskriterium og programmet ble modifisert for å tillate dette. Generaliseringsgraden i attributtrummet,  $G_A$ , ble representert ved størrelsen på vektoren mellom sammenslåtte enheter,  $\vec{v}$  fra Ligning 3, relativt til den totale variasjonen i datasettet som i Ligning 4

$$G_A = \frac{\vec{v}}{\sum_i^n \max_i - \min_i}$$

**Ligning 4**

der  $n$  er antall variable i datasettet. Utrykket under brøken tilsvarer vektoren mellom de to ytterpunktene i attributtrummet og kan beregnes på forhånd, det samme gjelder den absolutte lengden av  $\vec{v}$  ved en gitt generaliseringsgrad,  $G_A$ . Dermed trenger man bare å sjekke størrelsen på vektoren mellom sammenslåtte enheter etter hvert som algoritmen itererer og avbryte aggregeringen når denne når den gitte lengden. Har man nå to datasett med ulikt relieff medfører denne tilnærmingen at man kan sammenlikne avstanden mellom sammenslåtte enheter i de ulike datasettene, forutsatt at datasettene er normalisert/vektet med en felles transformasjon. Det betyr at man slipper å vurdere hvor stor generaliseringsgraden bør være for hvert enkelt datasett. Dersom relieffet er lavt vil man få en stor generalisering av enheter før vektoren når den ønskede lengden mens et datasett med høyere relieff vil bli avbrutt på et tidligere stadium.

### 3.3.2 Beregning av fasettparametere

Etter at datasettet er generalisert til et bestemt antall landskapsfasetter må nye parametere beregnes for hver av disse. Tilnærmingen gjort i dette studiet var å beregne gjennomsnittet av de 6 originale parametrene fra cellene innenfor hver enhet. Dette ble gjort for hver parameter med ArcInfo funksjonen "zonalmean" (ESRI, 2000) som tar et sonegrid (landskapsfasettene fra den kontekstuelle aggregeringen) og et verdigridd (en av de originale terrengparametrene) som input. Resultatet er et grid der alle celler i hver sone inneholder gjennomsnittet av parameterverdiene i den sonen. De resulterende gridene var dermed klare til bruk som variable i en klassifikasjon.

### 3.3.3 Klassifikasjon

Til selve klassifikasjonen ble det anvendt en ikke-styrt algoritme på et datasett med alle de 6 terrengparametrene som variable. Klasesentre ble definert ved hjelp av en iterativ clusteranalyse og deretter ble datasettet klassifisert med en maximum likelihood klassifikasjonsalgoritme. Begge algoritmene er implementert i ArcInfo som henholdsvis

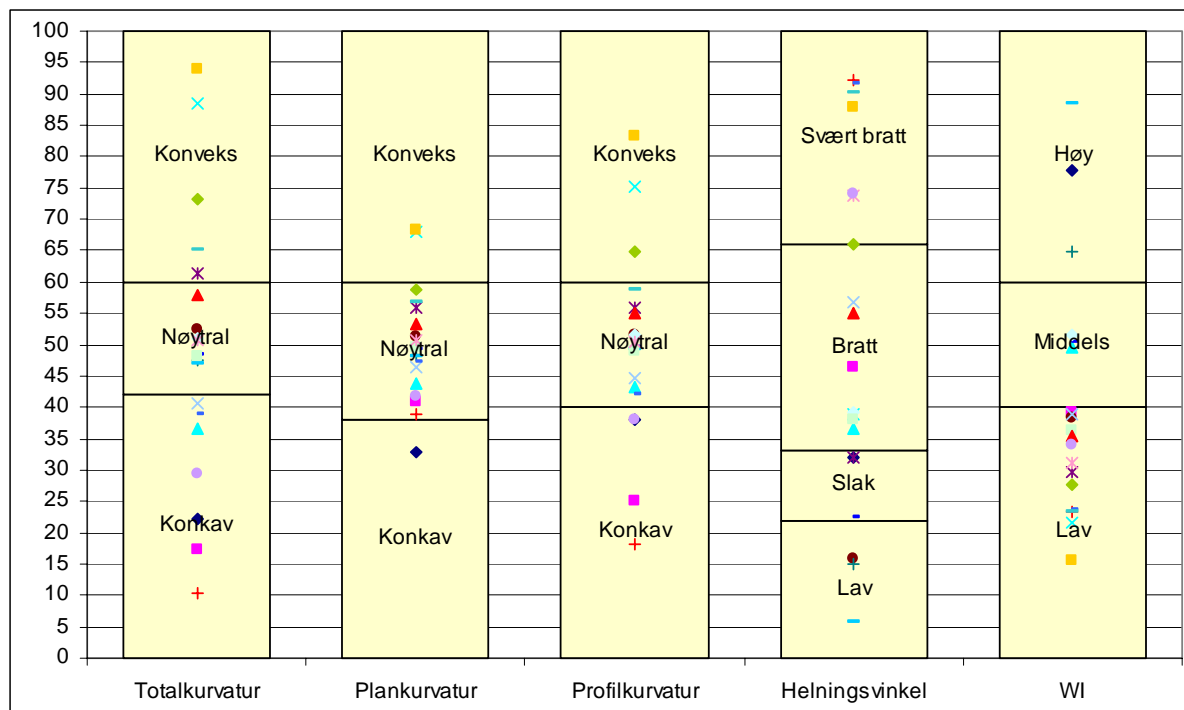


”isocluster” og ”mlclassify” (ESRI, 2000) og er beskrevet bl.a. i Burrough og McDonnell (1998). I clusteranalysen må man på forhånd bestemme antall resulterende klasser og maksimalt antall iterasjoner. Begge deler ble satt til 20 for alle analyser. Resultatet av analysen er en signaturfil som beskriver hver classes plassering og utbredelse i attributtrummet (middelverdi, varians og kovarians). Denne informasjonen brukes så for å tilegne hver celle i datasettet medlemskap til den klassen de har maksimum sannsynlighet (eng. *likelihood*) til. Resultatet av klassifikasjonen er et heltallsgrid med en klasseverdi for hver celle.

### 3.3.4 Kategorisering av landskapsklasser

Etter en ustyrt klassifikasjon ender man opp med klasser basert på datasettets fordeling i attributtrummet, vi kan kalle denne type klasser for clusterklasser. Vi er interessert i å vite hva slags topografisk region eller landformtype hver clusterklasse representerer så neste steg er å tilegne slik informasjon til klassene. Den mest opplagte måten å gjøre dette på er å sammenlikne forekomsten av clusterklasser med noen kjente informasjonsklassers forekomst i terrenget. Dette forutsetter imidlertid at det eksisterer referansedata fra studieområdet som man kan sammenlikne med. En annen tilnærming er å finne kjente informasjonsklassers plassering i attributtrummet og sammenlikne dette med clusterklassenes signatur. En slik analyse kan også gjøres uten faktiske felldata dersom man har en viss innsikt i hvordan informasjonsklassene karakteriseres av de gitte terrengparametrene, det er jo nettopp dette som menes med den geometriske signatur. For dette studiet var det ikke en prioritert oppgave å definere spesifikke geomorfologiske landskapsklasser tilstede i de ulike studieområdene, men det var viktig å kunne beskrive klassene godt nok til å kunne sammenlikne ulike klassifikasjoner med hverandre. Derfor ble hver klasse beskrevet ut fra sine topografiske egenskaper alene. Først og fremst var helningen ansett som en viktig faktor siden skråningsprosesser var i fokus. Klassene ble derfor kategorisert etter ”lav” (0°-10°), ”slak” (10°-15°), ”bratt” (15°-30°) eller ”svært bratt” (>30°) helning. På bakgrunn av Dikaus formler (Dikau, 1989) ble klassene også inndelt i kurvaturkategorier (konkav, konveks, plankonkav, plankonveks, profilkonkav, profilkonveks eller nøytral). Klassene fikk også et attributt for våthet dersom denne avvek fra det man skulle vente (jo slakere helning i klassen, jo høyere WI skulle til før klassen ble karakterisert som ”våt”). Resultatet var en tredelt beskrivelse av klassens helning, kurvatur og våthet. Vi ender da opp med kategorier som er beskrevet godt nok til at vi får et visst begrep om hva slags landskap det dreier seg om. Fordelen med denne måten å kategorisere på var at vi nå kun brukte topografiske mål for å

beskrive klassene slik at kategoriseringen kunne gjøres fullstendig objektivt, og dermed også automatisk.



**Figur 9.** Klassene ble kategorisert ved å analysere klassesentrenes fordeling innen hvert attributt (merk at enheter er normalisert fra 0-100). Figuren viser klassene fra Jotunheimen.

For å kunne oversette klassenes signatur fra kvantitative til kvalitative mål ble signaturene fra klassifikasjonen importert til et Excel regneark der grenseverdier kunne settes for hver kategori som i eksempelet i Figur 9. I ArcView kunne man nå bruke et Avenue script (makro programmeringsspråk for programvaren ArcView) til å lese klassenes signatur og grenseverdiene fra regnearket og bruke denne informasjonen til automatisk å lage en legende til klassifikasjonsresultatet. Dersom vi regner med at i hvert fall den *relative* sammenhengen mellom de ulike informasjonsklassene og terrengparametrene forblir noenlunde lik kan man nå sammenlikne klassifikasjonsresultater fra ulike relieff og skalaer.

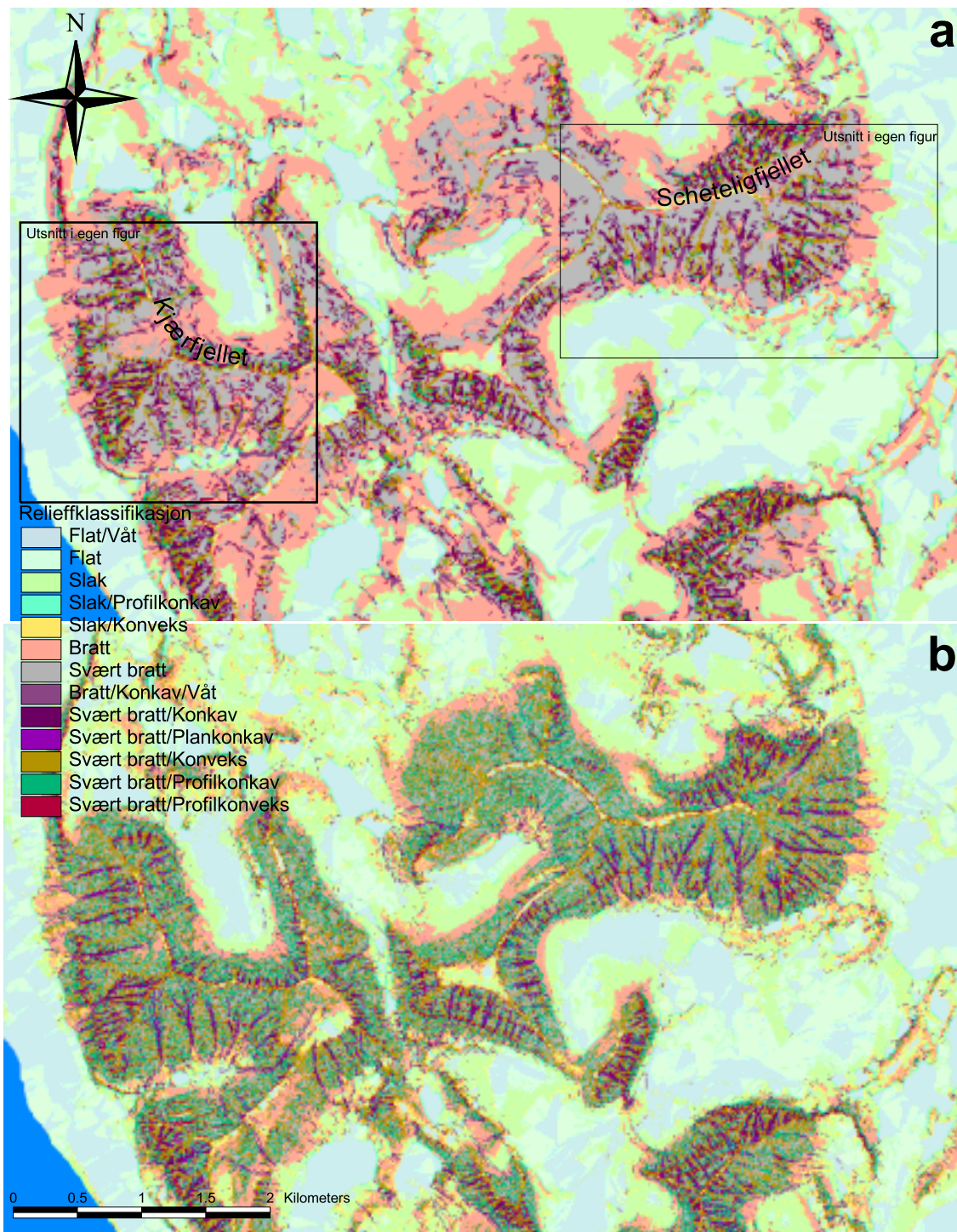
## 4 Resultater

### 4.1 Brøggerhalvøya

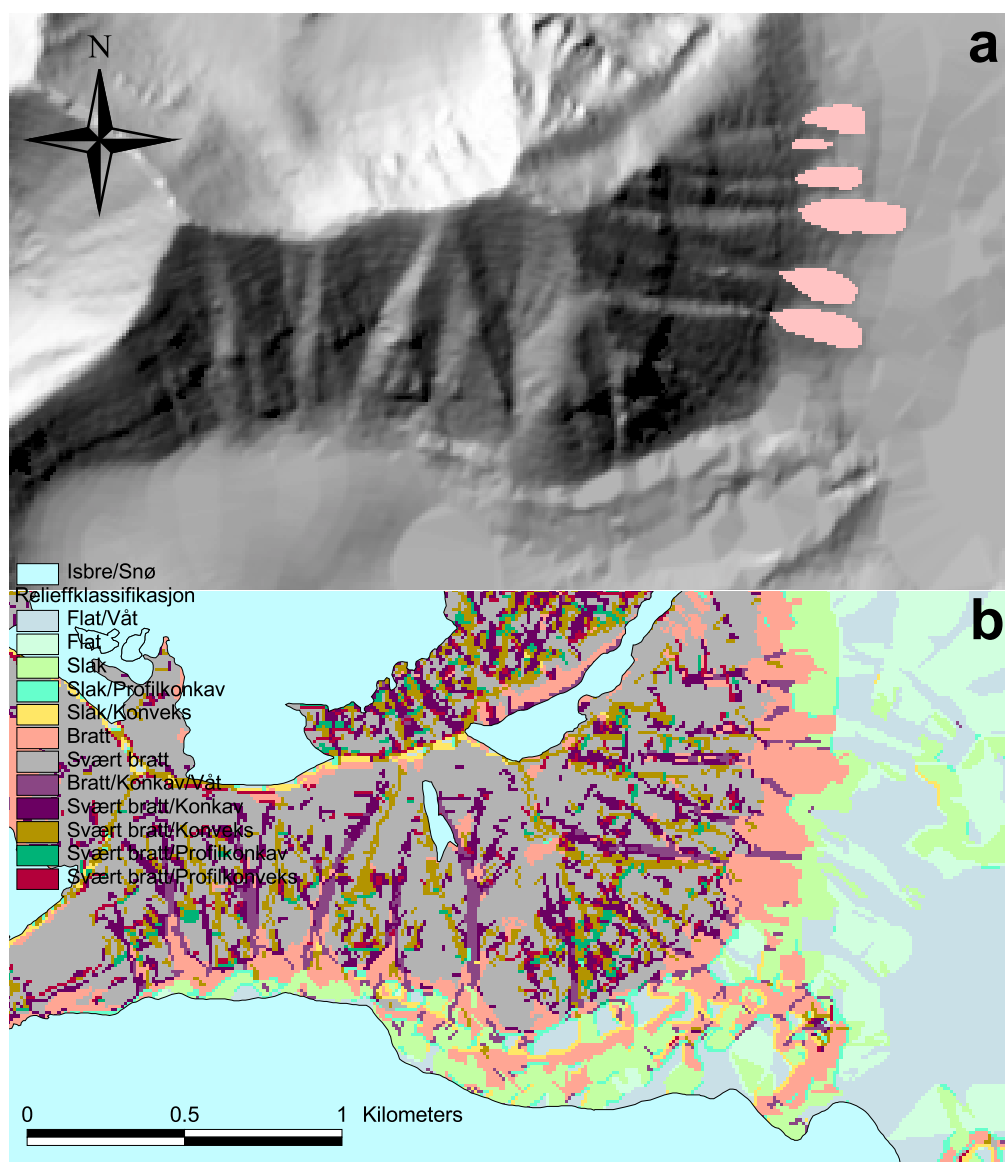
Formålet med studiet på Brøggerhalvøya var å kalibrere klassifikasjonsalgoritmene, sammenlikne den kontekstuelle klassifikasjonsmetoden med en ren isocluster klassifikasjon og evaluere klassifikasjonsresultatene. Dette var mulig fordi feltdata var tilgjengelig fra dette området i tillegg til en god terrengmodell.

Relieffklassifikasjonen beskrevet i avsnitt 3.3 ble anvendt på et 6,5×7,5 kilometer stort område som dekket hele den midtre delen av Brøggerhalvøya. Totalt relieff i området var på 710 meter. Generaliseringsgraden i attributtrommet for den kontekstuelle aggregeringen var på 15 % og tilsvarte en generalisering av datapunkter på omtrent 90 %. Etter klassifikasjonen kunne de 20 klassene plasseres i 13 ulike kategorier. For sammenlikning ble den samme klassifikasjonen anvendt på det samme området *uten* å først aggregere datasettet. Også denne klassifikasjonen resulterte i klasser som falt i de samme 13 kategoriene.

Klassene ble kategorisert etter helning, kurvatur og WI som beskrevet i avsnitt 3.3.4. Resultatet av de to klassifikasjonene kan ses i Figur 10a og b. Man ser at selv om det generelle mønsteret er likt i de to klassifikasjonene er klassene produsert via den kontekstuelle aggregeringen mer romlig konsistente og danner et mindre kaotisk bilde av landskapstypenes utbredelse. Utbredelsen av de kurvaturnøytrale, samt de fleste konkave klassene er noenlunde lik, men man ser at spesielt den profilkonkave og de konvekse klassene kompliserer bildet i den ikke-aggregerte klassifikasjonen.



**Figur 10. Relieffklassifikasjon av midtre Brøggerhalvøya med (a) og uten (b) kontekstuell aggregering**

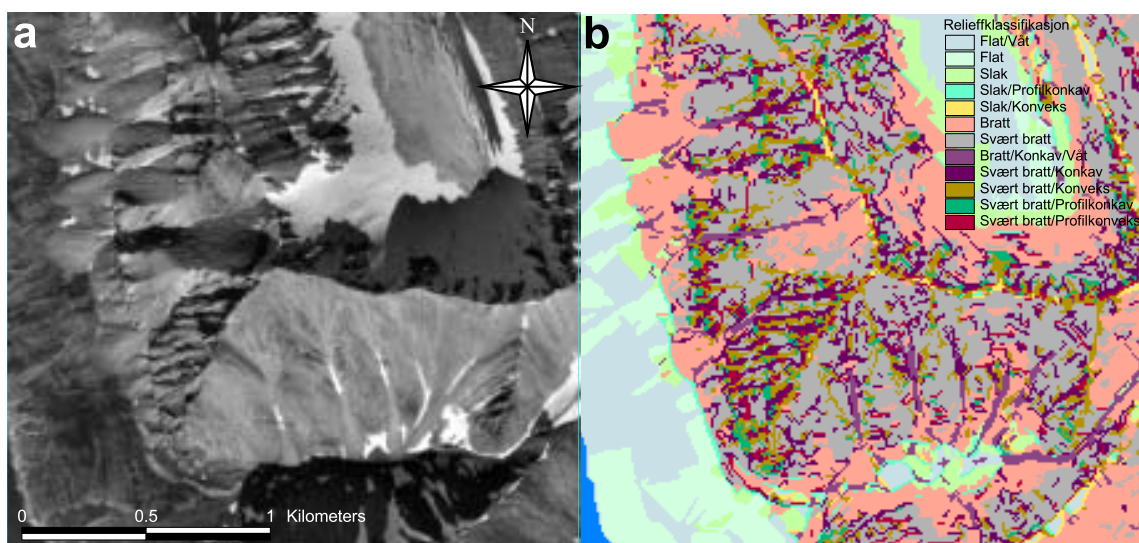


**Figur 11.** Skyggelagt terrengmodell over Scheteligfjellet med de kartlagte taluskjeglene (a) samt den kontekstuelle relieffklassifikasjonen (b) fra det samme området. Vi ser at taluskjeglene kjennetegnes av bratte skråninger umiddelbart nedenfor bratte/svært bratte konkave klasser (rasrenner). Under rasrennene på sydsiden av fjellet er det ikke dannet seg store taluskjeglere fordi løsmaterialet her er blitt transportert bort av breen.

I området rundt Scheteligfjellet (Figur 11) ser vi at de kartlagte taluskjeglene trer tydelig frem som forekomster av den bratte, rosa, klassen i overgangen mellom fjellskråningen og tundraen nedenfor. Denne klassen har ingen utpreget konveks plankurvatur slik man på grunn av talusenes kjegleform kanskje kunne forvente, i stedet ser det ut til å være forekomster av konkave klasser mellom talusene. Kartet i Figur 10 viser også at den rosa skråningsklassen som representerer selve talusformasjonene på ingen måte kan være ensbetydende med slike



former. Derimot ser det ut til at talusenes fordeling bedre forklares med kombinasjonen av denne klassen og forekomster av tilstrekkelig lange, bratte/svært bratte og konkave klasser (rasrenner) i overkant av skråningene (lilla). Fra Figur 10 ser vi at lange rasrenner forekommer på øst og sydsiden av Scheteligfjellet, samt på vest og sydsiden av Kjærfjellet. Sydskråningen av Scheteligfjellet ender på Brøggerbreen og store taluskjegler kan ikke ventes å forekomme her da breen vil transportere rasmaterialet vekk. På vest- og sydsiden av Kjærfjellet derimot finnes ingen bre og fra flybildet i Figur 12 kan man se at det finnes talusformasjoner også her.



**Figur 12.** Flybilde (a) og utsnitt av klassifikasjonen (b) fra vestsiden av Kjærfjellet.

Det er også andre typiske landformer som kommer frem i klassifikasjonsresultatene. Eggformasjoner trer frem som slake og konvekse klasser (gul) og rygger som bratte og konvekse klasser (brun). Den stratifiserte geologien i området gir seg ofte utslag i skrentformasjoner langs fjellsidene og disse formene trer frem som en kombinasjon av svært bratte profilkonvekse (mørk rød) og profilkonkave (mørk grønn) klasser. Alle disse mønstrene kommer tydeligere fram som mer homogene klasseforekomster i det kontekstuelle resultatet. I resultatet fra den rene isocluster klassifikasjonen trer mønstrene mer frem som tettere konsentrasjoner av klasseforekomster som ellers finnes spredd ut over de bratte områdene.

Det var mulig å gjøre en kvantitativ vurdering og sammenlikning av de to klassifikasjonsresultatene basert på felldataene. Dette ble gjort ved å sette opp klasseforvirringsmatriser (eng. *class confusion matrix*) (se Tabell 1, Tabell 2 og Tabell 3). For

hver klassifisert klasse settes det da opp antall forekomster av celler fra de ulike referanseklassene. I dette tilfellet var det egentlig bare *en* referanseklasse, nemlig de kartlagte skråningsformene, i tillegg til at en del celler som helt sikkert *ikke* tilhørte denne klassen var kjent (hhv ”talus” og ”annet” i tabellene). De klassene som representerte taluser kunne lett identifiseres ved å sammenholde klassifikasjonsresultatet med referansekartet, resten av klassene ble kategorisert som ”annet”. Ved å summere alle korrekt klassifiserte celler (diagonalen i forvirringsmatrisen) og dividere dette på totalt antall celler får man et enkelt, men udetaljert mål på total nøyaktighet. Man kan også beregne det som kalles produsents og brukers nøyaktighet til hver klasse (eng. *producer's accuracy/user's accuracy*, eventuelt *errors of omission/errors of commission*) (Congalton, 1991). Produsentnøyaktigheten til en klasse beregnes ved å dividere antallet celler i klassen som har blitt korrekt klassifisert med det totale antallet celler i referanseklassen. Størrelsen sier noe om sannsynligheten for at referanseklassen er blitt korrekt klassifisert og blir kalt produsents nøyaktighet fordi det fra en produsents synspunkt vil være interessant å vite hvor stor andel av en gitt informasjonsklasse som kan klassifiseres. Brukers nøyaktighet for en klasse kalkuleres ved å dele antallet korrekt klassifiserte celler med det totale antallet celler i klassen og uttrykker sannsynligheten for at akkurat denne klassen faktisk representerer referanseklassen. Størrelsen har fått navnet brukers nøyaktighet fordi det fra en brukers synspunkt er mest interessant å vite hvor sannsynlig det er at det han ser på kartet faktisk forekommer på bakken.

**Tabell 1. Klasseforvirringsmatrise med produsents og brukers nøyaktighet for det kontekstuelle klassifikasjonsresultatet.**

Klassifikasjon	Referansedata		Evaluering	
	Talus	Annet	Produsents nøyaktighet	Brukers nøyaktighet
Talus	729	302	75 %	71 %
Annet	247	2110	87 %	90 %
Sum	976	2412		
Total nøyaktighet				84 %

**Tabell 2. Forvirringsmatrise for ikke-aggregert klassifikasjon.**

Klassifikasjon	Referansedata		Evaluering	
	Talus	Annet	Produsents nøyaktighet	Brukers nøyaktighet
Talus	647	404	66 %	62 %
Annet	326	2008	83 %	86 %
Sum	976	2412		
Total nøyaktighet				78 %

**Tabell 3. Forvirringsmatrise for ikke-aggregert klassifikasjon etterbehandlet med et majoritetsfilter.**

Klassifikasjon	Referansedata		Evaluering	
	Talus	Annet	Produsents nøyaktighet	Brukers nøyaktighet
Talus	587	379	68 %	61 %
Annet	280	1682	82 %	86 %
Sum	867	2061		
Total nøyaktighet				77 %

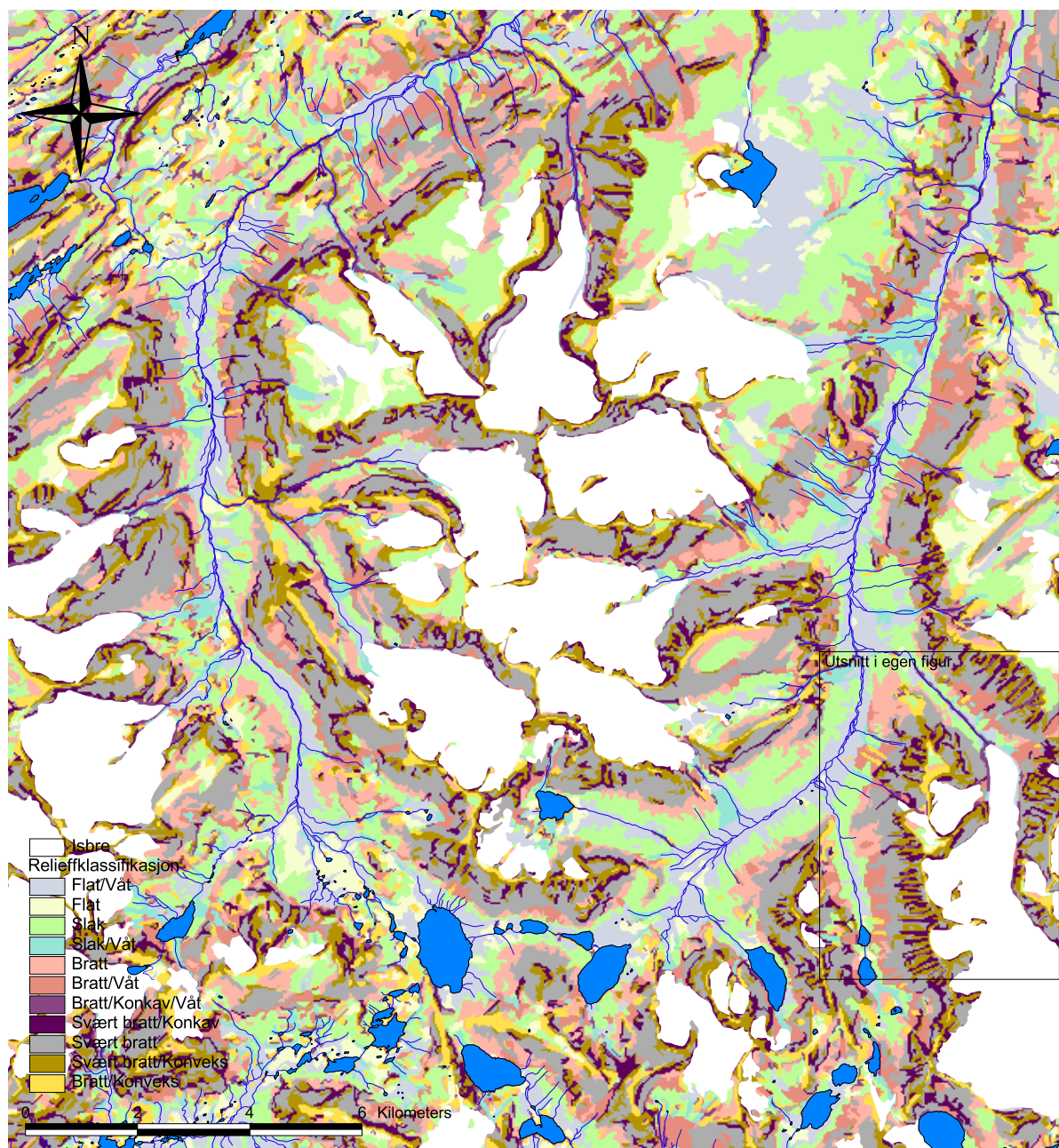
Vi ser av Tabell 1 og Tabell 2 at den kontekstuelle klassifikasjonen faktisk oppnådde et bedre resultat for alle evalueringsmål, men man må huske at referansedataene er begrenset til et lite område og at bare en virkelig klasse er representert. Referansedataene oppfyller altså ikke kravene som stilles for å gjøre en god nøyaktighetsvurdering uten bias.

Tabell 3 viser forvirringsmatrisen til det ikke-aggregerte klassifikasjonsresultatet etter å ha anvendt et 3×3 celler stort majoritetsfilter på det. Dette er en vanlig måte å etterbehandle en klassifikasjon av romlige data på da det fjerner små isolerte klasseforekomster ved at hver celleverdi erstattes med verdien den av dens naboer det finnes mest av. Vi kan se av tabellen at denne operasjonen hadde svært liten påvirkning på nøyaktigheten av klassifikasjonen (forskjellen i totalt antall evaluerte celler mellom Tabell 2 og Tabell 3 skyldes at filtreringen medfører celler uten verdi langs kanten av datasettet).

## 4.2 Jotunheimen

Terrengmodellen fra Jotunheimen hadde en oppløsning på 30×30 meter og et totalt relieff på 1970 meter. På denne modellen ble en kontekstuell relieffklassifikasjon, tilsvarende den utført på Brøggerhalvøya, anvendt. Målet var å teste hvordan klassifikasjonsalgoritmen fungerte på et annet relieff og en terrengmodell med lavere oppløsning. Det var også interessant å se om en kategorisering av klassene etter de samme topografiske kriteriene som ble brukt på Brøggerhalvøya ga meningsfulle klassebeskrivelser. Før aggregeringen ble de samme fire parametrene normalisert og vektet som på Brøggerhalvøya og deretter ble de resulterende fasettene klassifisert ved hjelp av de samme seks attributtene som før. Generaliseringsgraden i attributtrummet på 15 % tilsvarte her bare omtrent 80 % generalisering av datapunkter. Etter at klassifikasjonen var gjennomført ble signaturfilene analysert som beskrevet i avsnitt 3.3.4 og hver klasse ble plassert i sin respektive kategori. De 20 originale klassene endte da opp i 11 ulike kategorier. Deretter kunne resultatet evalueres visuelt ved hjelp av grunnlagskart med høydekoter og elver (N50, kartblad 15182, "Galdhøpiggen").

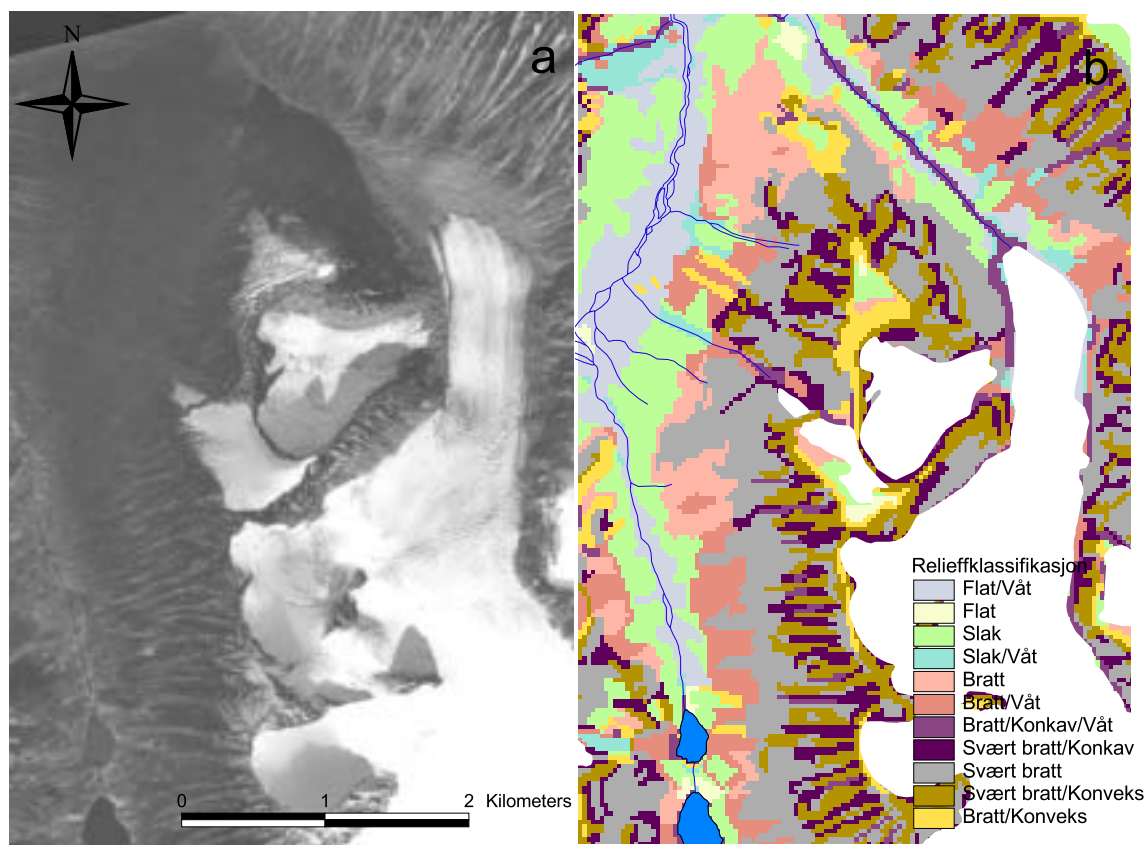




**Figur 13. Kontekstuell klassifikasjon av 30×30 meters modellen over Jotunheimen**

Når vi ser på klassifikasjonen av Jotunheimen (Figur 13) ser vi at både selve klassifikasjonen samt kategorisering av klassene så ut til å ha produsert mer eller mindre fornuftige resultater. Vi kan kjenne igjen former som egger og rygger fra den forrige analysen, men eggene har her fått en brattere helning enn vi hadde på Brøggerhalvøya. En viss forskjell er det også blant de konkave klassene. Mens disse stort sett forekom langs skråningsretningen (rasrenne/ravine) på Brøggerhalvøya, finnes de her også i større grad normalt på skråningene (skrent). Fra

analysen av Brøggerhalvøya fant vi at forekomsten av raskjegler gjerne kunne forklares med forekomster av lange bratte og konkave klasser med mindre bratte skråningsklasser i underkant. Det er spesielt området rundt Heilstugubreen som i denne klassifikasjonen utmerker seg for et slikt mønster. Undersøkelse av flybilder viste at det også her virkelig var raskjegler (se Figur 14).



**Figur 14. Flybilde (a) og utsnitt fra klassifikasjonsresultatet (b) ved Heilstugubreen.**

## 5 Diskusjon

### 5.1 Sammenlikning av metodene

Den kontekstuelle aggregeringen av datasettet skaper en romlig koherens i klassifikasjonsresultatet som mangler ved en ren celle for celle klassifikasjon, men ellers er det generelle mønsteret i de to klassifikasjonene svært likt. Dette tyder på at aggregeringen av datasettet, hvis en fornuftig generaliseringsgrad er valgt, ikke kompromitterer for mye av den originale informasjonen i datasettet. Dermed blir resultatet av klassifikasjonen mer eller mindre det samme enten datasettet er aggregert eller ikke, men den aggregerte klassifikasjonen skaper klasser som er lettere å tolke visuelt. De romlig koherente klasseforekomstene er også lettere å bruke direkte i videre analyse. Etter en vanlig klassifikasjon ville man for eksempel måtte gjort en majoritetsfiltrering for å fjerne prikkmønstrede klasseforekomster som ofte inntreffer der en har en viss tekstur i terrengmodellen. Slike klasseforekomster har aggregeringsalgoritmen for en stor del allerede slått sammen med en nabo, med den åpenbart positive forskjell at utliggerer ikke automatisk slås sammen med den av naboklassene den har størst grenseflate til, men med den som er likest. Vi så jo også fra Tabell 2 og Tabell 3 at en slik operasjon hadde liten eller ingen påvirkning på nøyaktigheten av klassifikasjonsresultatet.

Evalueringen av klassifikasjonsresultatet viste også at den kontekstuelle klassifikasjonen skapte klasser med bedre nøyaktighet. Referansedataene brukt i denne evalueringen var imidlertid av en slik art at man ikke av dette kan konkludere med at den kontekstuelle klassifikasjonen produserte et mer nøyaktig resultat. Likevel forteller det oss at aggregeringen ikke nødvendigvis er et kompromiss mellom nøyaktighet og lesbarhet/anvendbarhet. Man kan også argumentere for at resultatet av en kontekstuell klassifikasjon burde produsere et mer nøyaktig resultat fordi landformer i første omgang er romlige enheter. Krav til homogenitet i rommet burde altså overstyre kravet til homogenitet til geometrien. Videre er det alltid slik at en opererer på en skala. ”Hvor lang er den britiske kyst?” spurte Mandelbrot (1967) og påpekte at dette spørsmålet ikke hadde ett svar, men var bestemt av hvor nøyaktig du klarte å måle, hvilken skala du opererte på. Nå var ikke dette i seg selv Mandelbrots hovedbudskap, men det er interessant som eksempel. Den britiske kyst er, som en landform og alle fenomener med en utstrekning for øvrig, sammensatt av et hierarki av enheter. I en gridbasert terrengmodell er den minste enheten en gridcelle og man kan ikke måle objekter som er mindre enn dette, Pike (1988) definerer de minste landformer som kan registreres i en

terrengmodell til å være omtrent dobbelt så store som cellestørrelsen. Det originale datasettet forteller oss om forholdene i hver av enhetene og en direkte attributtbasert klassifikasjon av disse forenkler per definisjon kun informasjonen i attributtene. Romlig sett er altså informasjonen fortsatt på cellenivå. Nå er det i midlertidig slik at landformene som ble registrert som referansedata er sammensatt av flere celler (stort sett mellom 100 og 300) og det var ikke registrert noen variasjon innenfor hver av disse. Vi er altså egentlig ikke interessert i forholdene i hver celle, men i forholdene i hver landform og det er dette vi forsøker å fremskaffe med den kontekstuelle aggregeringen. Sammenlikner vi resultatkartene i Figur 10 ser vi at mens raskjeglene stort sett er representert med kun *en* klasse i det kontekstuelle resultatet er de en kombinasjon av en bratt klasse og en slak/konveks klasse i det andre resultatet. Dette er jo ikke nødvendigvis en gal klassifikasjon, sannsynligvis er det en bedre beskrivelse av taluskjeglen fordi den nettopp består av en delvis plankonveks skråning. Men dette er en grad av nøyaktighet som er høyere enn det vi ønsker å måle og som er på en annen skala enn det referansedataene er samlet inn på. Dermed kommer også klassifikasjonsresultatet dårligere ut i nøyaktighetsevalueringen.

## **5.2 Romlig modellering av landformer og prosesser**

For deteksjon av områder med raskjeglør så klassifikasjonsmetoden ut til å være egnet. Dette kan vi si fordi flybilder viste at det faktisk var slike former der en kunne forvente det ut i fra klassifikasjonsresultatet både på Brøggerhalvøya og i Jotunheimen. Resultatene viser altså at klassifikasjonen av de utvalgte parametrene skaper topografiske regioner som gir mening fra et geomorfologisk synspunkt. For å forklare fordelingen av taluskjeglør kunne vi bruke kombinasjonen av to ulike topografiske regioner. Selve raskjeglen ligger på en topografisk klasse med bratt helning, men ikke brattere en ca.  $30^\circ$ . I tillegg måtte det i overkant av denne skråningen ligge en bratt/svært bratt konkav region med stor intern høydeforskjell. Denne klassen representerer rasrenner som leder forvitningsmateriale ned til raskjeglen. Siden regionene er dannet ved klassifikasjon av morfometriske parametre kan sammenhengen mellom disse og landformene også uttrykkes som morfometriske betingelser for fordelingen av den utvalgte landformtypen. Studiet tyder på at disse betingelsene har en generell karakter i og med at de så ut til å være gyldige i klassifikasjonsresultater fra to terrengmodeller med både ulikt relieff og skala.

De morfometriske betingelsene trenger ikke i seg selv å relateres direkte til skråningsprosesser, men vi kan se at spesielt betingelsene knyttet til helningsvinkel kan

baseres på fysiske prosesser. En raskjegle vil for eksempel ikke kunne bygge seg opp til en helning stort brattere en  $30^\circ$  fordi friksjonen mellom fragmentene i skråningen da ikke lenger ville være stor nok til å hindre løsmaterialet å rase ut å stabilisere seg ved en lavere helningsvinkel. Videre så det også ut til å være en sammenheng mellom rasrennenes størrelse og størrelsen til den nedenforliggende taluskjeglen. Selv om det siste i og for seg ikke burde være noen overraskelse gir det en klar indikasjon på at den romlige sammenhengen mellom de to klassene bygger på fysiske prosesser styrt av gravitasjonskraften og at denne sammenhengen muligens kan kvantifiseres.

Selv om modellen er basert på empiriske relasjoner og ikke modellerer de faktiske fysiske prosessene i skråningene kan de morfometriske betingelsene brukt i modellen være til hjelp for å definere den romlige variabiliteten til ulike landformer og dermed også prosesser. Slik modellering kan ha betydning for mange applikasjoner innen geomorfologi. Et konkret eksempel vil kunne være å knytte forvittringsrater til raskjeglenes størrelse for å bruke dette til en regional modellering av landskapsevolusjon og historisk isbreutbredelse. På Brøggerhalvøya så vi hvordan rasrenner som ledet ned mot breer ikke produserte raskjegler fordi breen transporterer forvittringsmaterialet bort. Dersom vi ser for oss at områder der det nå er raskjegler tidligere har vært nediset vil vi kunne anta tidspunktet for breens tilbakesmelting ved å beregne taluskjeglens alder. Dette vil være en spesielt interessant problemstilling på Svalbard da det er stor usikkerhet rundt utbredelsen av breer i dette området under siste istid.

### **5.3 Skala**

Studiet viste at metoden kunne brukes uforandret i ulike relieff og på terrengmodeller med ulik oppløsning og likevel produsere et tilfredsstillende resultat. Dette tyder på at de topografiske betingelsene som brukes til å modellere taluskjeglenes utbredelse er skalauavhengige innenfor den målestokken vi her har jobbet på. Det ble gjort noen tester på å resample terrengmodellen fra Jotunheimen til en lavere oppløsning og det ble oppdaget at når celledstørrelsen nærmet seg  $100 \times 100$  meter var det ikke lenger tydelig i klassifikasjonsresultatet hvor taluskjegler burde forekomme. Dette skyldes naturlig nok at klassene som indikerer rasaktivitet rett og slett ikke er store nok til å bli representert i terrengmodellen, med andre ord beveger terrengmodellens oppløsning seg utenfor skalaen til det vi søker å observere. Klassen som representerer eggformasjoner i Jotunheimen er brattere enn de var på Brøggerhalvøya, dette er nok et resultat av oppløsningen på terrengmodellen

som gjorde at klassifikasjonen ikke klarte å skille mellom den konsentrerte konkaviteten langs en egg og de bratte skråningene som forekommer umiddelbart på hver side. I Jotunheimen klarte klassifikasjonen i mindre grad å skille mellom bratte plankonkave klasser som representerte raviner, og de som representerte skrenter. Om dette er et resultat av oppløsningen eller ulike geomorfologiske og/eller geologiske betingelser som gjør at det faktisk finnes flere plankonkave skrenter i Jotunheimen er vanskelig å si, men resultatet var i hvert fall at lokaliseringen av taluskjegler måtte gjøres semimanuelt i og med at klasseforekomstenes lengderetning i forhold til helningsretningen måtte undersøkes. Dette er ingen vanskelig oppgave, men problemet kunne kanskje vært unngått ved for eksempel å inkludere høydeforskjellen innen hver av de aggregerte landskapsfasettene. Dette er et tema som diskuteres videre i neste avsnitt

#### **5.4 *Utnyttelse av høyere ordens parametere***

Enhetene generert fra aggregeringen kan karakteriseres som homogene topografiske regioner. Hvis vi antar at valg av parametere og grad av generalisering var noenlunde hensiktsmessig vil disse enhetene også kunne representere faktiske landformelementer med relativt liten variasjon i geometriske attributter. Man kan altså anta at hvert element representerer en forekomst av et landformelement og at det i hvert element er gitte fysiske prosesser som er dominerende. I denne oppgaven ble imidlertid ikke det fulle potensialet som ligger dette utnyttet. Når elementets attributter, eller fasettparametrene, ble beregnet ble dette gjort ved å ta gjennomsnittet av parametere først beregnet for hver celle innen elementet. Fordelen med denne tilnærmingen var at den var enkel og rask samt at det da ble mer hensiktsmessig å sammenlikne den aggregerte klassifikasjonen med den ikke-aggregerte. Prosessene som virker i en topografisk region er imidlertid ikke bare avhengig av prosessene i hvert punkt innen regionen, men også av regionens form, størrelse og interne topografi. Et eksempel kan være de konkave, bratte klassene som, spesielt i Jotunheimen forekom som avlange enheter både langs og normalt på skråningsretningen. Cellene innen klasseforekomstene hadde hver for seg nokså like egenskaper, men man kan lett forestille seg at prosessene som virker i de to ulike regionene vil være vidt forskjellige. McDermid og Franklin (1995) demonstrerer bruken av geomorfometriske parametere beregnet for skråningsnivå, i kombinasjon med parametere for hvert punkt/celle i skråningen. Dette forbedret resultatet av en landskapsklassifisering betraktelig framfor å bruke kun en av parametertypene og det meste av forbedringen kunne forklares med påvirkningen fra de skråningsbaserte parametrene.

Resultatene viser at for å få til en fullstendig topografisk beskrivelse av de utvalgte landformene bør man bruke både landformens geometriske egenskaper (geometrisk signatur), dens interne og *eksterne* topologi. Dette påpekes også bl.a. av Schmidt og Dikau (1999), men analyse av romlige sammenhenger mellom denne type objekter kan være en komplisert affære som krever svært avanserte algoritmer. Dette fordi man i motsetning til regulære romlige analyser i en gridstruktur, ikke arbeider med jevnt fordelte punkter, men med objekter av ulik størrelse og form.

### **5.5 Valg av terrengparametere**

Terrengparametrene som ble brukt som variable både i aggregeringen og klassifikasjonen av datasettet var i utgangspunktet utpekt på bakgrunn av deres antatte betydning for formene/prosessene i fokus. I en klassifikasjonsprosess bør de ulike variablene i så liten grad som mulig og det ble derfor først gjort en del forsøk på klassifikasjon der de mest korrelerte variablene ble utelatt (totalkurvaturen som hadde høy korrelasjon med både plan- og profilkurvatur samt plankurvaturen som korrelerte med WIplan). Når resultatet av ulike klassifikasjoner ble sammenliknet med referansedataene viste det seg at både taluskjegler og rasrenner ble bedre avgrenset dersom alle variablene var inkludert i datasettet. Årsaken til at bruken av korrelerte variable så ut til å gi et bedre resultat var sannsynligvis at dette medfører en implisitt vektning av de aktuelle variablene i klassifikasjonen. Siden hovedpoenget med dette studiet var å vise hvordan klassifikasjonsresultatet kunne forbedres med den kontekstuelle algoritmen ble ikke valget og bruken av parametere tillagt så mye vekt, men skulle metodene videreutvikles slik at de kunne produsere en mer helhetlig landskapsklassifikasjon er valget av terrengparametere en svært viktig del av prosessen som bør vurderes nøye slik at de reflekterer kjente egenskaper til fenomenene som studeres (Weibel og DeLotto, 1988).

I mange terrengklassifikasjoner er både høyden selv og helningsretningen (eller parametere avledet av denne) ofte brukte variable. Den viktigste grunnen til at disse ikke ble brukt i dette studiet var at begge parametrene har stor romlig avhengighet. Det vil si at bruken av disse automatisk introduserer en romlig koherens i klassifikasjonsresultatet. Det ville dermed blitt vanskeligere å skille hvor stor del av den romlige koherensen som var et resultat av romlig avhengige variable eller aggregeringsalgoritmen. Siden studiet også tok sikte på å konsentrere seg om visse skråningsforer, og en helhetlig beskrivelse av alle klassene i resultatet ikke ble gitt, var i denne sammenhengen uinteressant om en skråning vendte mot syd eller nord etc.

Alternativt kunne nok i hvert fall helningsretningen blitt brukt i aggregeringsalgoritmen for å tilnærme seg for eksempel Dikaus landskapsfasetter (Dikau, 1989), men selv om dette kanskje ville forbedret klassifikasjonsresultatet ville det da altså vært mindre hensiktsmessig å sammenlikne det aggregerte med det ikke-aggregerte resultatet.

## **5.6 Generaliseringsgrad**

Avbruddskriteriet på 15 % generalisering i attributtrommet medførte en mye større generalisering av datapunkter i datasettet fra Brøggerhalvøya enn det fra Jotunheimen. Dette kan forventes ettersom det totale relieffet er høyere i Jotunheimen og fordi en større del av terrenget på Brøggerhalvøya var mer eller mindre flatt, og aggregeringen av disse områdene førte i liten grad til generalisering i attributtrommet. Selve målet for generaliseringsgrad,  $G_A$  fra Ligning 4, var et nokså enkelt mål som kun tok hensyn til ytterpunktene i datasettet og ikke fordelingen av data mellom disse. Det vil si at målet kun er meningsfullt så lenge dataene er mer eller mindre normalfordelt og ikke har ekstreme verdier. Videre sier dette generaliseringsmålet ingen ting om hvilke variable generaliseringen forekommer i. Dersom for eksempel alle variable bortsett fra *en* er like for to naboenheter vil forskjellen mellom enhetene i denne ene variabelen kunne være svært stor uten at avstandsvektoren overstiger maksimal lengden. En annen tilnærming til avbruddskriteriet kunne være å definere maksimal generalisering, eller forskjell, for hver variabel i stedet for, eller i tillegg til, maksimal total generalisering. På denne måten ville det være lettere å definere avbruddskriterier for ulike typer terreng, skala og applikasjoner ved bruk av feltdata eller generell kunnskap til landformene og attributtene som analyseres.

## **5.7 Klassifikasjonsalgoritmer**

I dette studiet ble mye arbeid lagt i å se på effekten av ulike variabelkombinasjoner, transformasjoner og generaliseringsgrader i aggregeringsalgoritmen. Effekten av ulike antall klasser og ulike algoritmer i selve klassifikasjonen ble derimot ikke undersøkt. En kontinuerlig klassifikasjonsalgoritme, tilsvarende de beskrevet av for eksempel Irvin et al. (1997) eller Burrough et al. (2000), kunne gitt opplysninger om variasjonen innen og mellom klassene og ville kunnet tilby informasjon om usikkerheten innen hver klasseforekomst. De siste årene er det også blitt rettet oppmerksomhet mot bruken av neurale nettverk for å inkorporere romlig eller temporær kontekst inn i analyse av satellittbilder (Lees, 1996). Dette er algoritmer som "læres opp" til å ta avgjørelser basert på tidligere "erfaringer". Anvendelsen av slike til automatisk landskapsklassifikasjon bør undersøkes i videre studier.



## 6 Konklusjon

Studiet har vist at en kontekstuell klassifikasjon, gjennomført ved en preliminær aggregering av et datasett, produserte klasser som var vel så geomorfologisk relevante som de produsert ved en ren attributtbasert klassifikasjon. Den kontekstuelle klassifikasjonen skapte også et mer romlig koherent resultat som var både lettere å tolke visuelt og å bruke direkte i videre analyser. Metoden var i stand til å oppdage områder der raskjegler forekommer i to ulike datasett med ulikt relieff og ulik oppløsning. Dette underbygger antagelsen om at det er en sammenheng mellom topografiske parametere, landformer og prosesser, og at slike sammenhenger, dersom de kvantifiseres, kan brukes til romlig modellering av geomorfologiske prosesser. Resultatene indikerer også en klar og muligens kvantifiserbar, sammenheng mellom ulike landformsklasser. Dette fordi forekomsten av taluskjegler i dette studiet kunne forklares med kombinasjonen av rasrenneklasse i overkant av middels bratte skråningsklasser, og at taluskjeglens størrelse så ut til å korrelere med størrelsen på raserennene.

Det antydes altså at den kontekstuelle aggregeringen definerer enheter som kan knyttes direkte til ulike landformer. Dette faktum burde utnyttes og videre studier burde bl.a. se på effekten av å bruke parametere beregnet for disse enhetene som tar hensyn til deres interne og eksterne topologi. Først da kan det fulle potensialet som ligger i aggregeringen utnyttes. Videre burde også metoder utvikles for kvantitativ analyse av interklasse-sammenhenger, det vil si den romlige fordelingen av ulike klasseforekomster i forhold til hverandre på den tredimensjonale terrengoverflaten.

## Referanser

- Beven, K. J. og Kirkby, M. J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences*, 24: 43-69.
- Beven, K. J., Kirkby, M. J., Schofield, N. og Tagg, A. F., 1984. Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments. *Journal of Hydrology*, 69(1-4): 119-143.
- Bjørnstad, A. E., Bryhni, T. S. og Skaug, Ø., 1998. *Prosjektoppgave GG372*, Geografisk institutt, Universitet i Oslo, Oslo.
- Budd, W. F. og Warner, R. C., 1996. A computer scheme for rapid calculations of balance-flux distributions. *Journal of Glaciology*, 23: 21-27.
- Burrough, P. A., 1986. Principles of geographical information systems for land resources assessment. Monographs on soil and resources survey, 12. Clarendon Press, Oxford, 194 pp.
- Burrough, P. A. og McDonnell, R. A., 1998. Principles of geographical information systems. Oxford University Press, Oxford.
- Burrough, P. A., van Gaans, P. F. M. og MacMillan, R. A., 2000. High-resolution landform classification using fuzzy k-means. *Fuzzy sets and systems*, 113(1): 37-52.
- Campbell, J. B., 1996. Introduction to remote sensing. Taylor & Francis, London.
- Congalton, R. G., 1991. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment*, 37: 35-46.
- Dikau, R., 1989. The application of a digital relief model to landform analysis in geomorphology. In: J. Raper (Editor), *Three dimensional applications in geographical information systems*. Taylor and Francis, London, United Kingdom, pp. 51-77.
- ESRI, 2000. ArcInfo Online Manual. Environmental Systems Research Institute (ESRI), Inc., Redlands, CA, USA.
- Etzelmüller, B., Ødegård, R., Berthling, I. og Sollid, J. L., 2001. Terrain parameters and remote sensing data in periglacial reseach. *Permafrost and Periglacial Processes*, 12(1): 79-92.
- Evans, I. S., 1972. General geomorphometry, derivatives of altitude, and descriptive statistics. In: R. J. Chorley (Editor), *Spatial analysis in geomorphology*. Mathuen & Co Ltd, pp. 17-92.
- Evans, I. S., 1998. What do terrain statistics really mean? In: S. N. Lane, K. S. Richards og J. H. Chandler (Editors), *Landform Monitoring, Modelling and Analysis*. Wiley, Chichester, pp. 120-138.

- Friedrich, K., 1996. Multivariate distance methods for geomorphographic relief classification. In: H. J. Heineke et al., *Proceedings EU Workshop on Land Information Systems: Developments for planning the sustainable use of land resources*, Hannover. European Soil Bureau, pp. 259-266.
- Hutchinson, M. F., 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology*, 106(3-4): 211-232.
- Irvin, B. J., Ventura, S. J. og Slater, B. K., 1997. Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. *Geoderma*, 77(2-4): 137-154.
- Lees, B. G., 1996. Neural network applications in the geosciences: an introduction. *Computers and Geosciences*, 22(9): 955-957.
- Mandelbrot, B., 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimensions. *Science*, 156(3775): 636-638.
- Mark, D. M., 1975. Geomorphometric parameters: a review and evaluation. *Geografiska Annaler*, 57A: 165-177.
- McDermid, G. J. og Franklin, S. E., 1995. Remote sensing and geomorphometric discrimination of slope processes. *Zeitschrift für Geomorphologie, Supplementband*, 101: 165-185.
- Moore, I. D., Gessler, P. E., Nielsen, G. A. og Peterson, G. A., 1993a. Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, 57(2): 443-452.
- Moore, I. D., Grayson, R. B. og Ladson, A. R., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5(1): 3-30.
- Moore, I. D., Turner, A. K., Wilson, J. P., Jenson, S. K. og Band, L. E., 1993b. GIS and land-surface-subsurface process modeling. In: M. F. Goodchild, B. O. Parks og L. T. Steyaert (Editors), *Environmental modeling with GIS*. Oxford University Press, New York, NY, pp. 196-230.
- Pike, R. J., 1988. The geometric signature; quantifying landslide-terrain types from digital elevation models. *Mathematical Geology*, 20(5): 491-511.
- Pike, R. J., 1995. Geomorphometry - progress, practice, and prospect. *Zeitschrift für Geomorphologie, Supplementband*, 101: 221-238.
- Pike, R. J., 2000. Geomorphometry - diversity in quantitative surface analysis. *Progress in Physical Geography*, 24(1): 1-20.
- Schmidt, J. og Dikau, R., 1999. Extracting geomorphometric attributes and objects from digital elevation models - semantics, methods, future needs. In: R. Dikau og H. Saurer (Editors), *GIS for Earth surface systems. Analysis and modeling of the natural environment*. Gebrüder Borntraeger, Germany, pp. 154-171.

- Sulebak, J. R., 1999. Fractal analysis of surface topography. *Norsk geografisk tidsskrift*, 53(4): 213-225.
- Sulebak, J. R., Etzelmuller, B. og Sollid, J. L., 1997. Landscape regionalization by automatic classification of landform elements. *Norsk Geografisk Tidsskrift*, 51(1): 35-45.
- Weibel, R. og DeLotto, J. S., 1988. Automated terrain classification for GIS modeling, *GIS/LIS '88. Proceedings from the 3rd international conference*, San Antonio. ASPRS/ACSM, pp. 618-627.
- Zevenbergen, L. W. og Thorne, C. R., 1987. Quantitative-analysis of land surface-topography. *Earth Surface Processes and Landforms*, 12(1): 47-56.

## **Vedlegg: Improving Relief Classification with Contextual Merging**

Artikkel publisert i Bjørke, Jan Terje and Håvard Tveite (eds.), *Proceedings of the 8th Scandinavian Research Conference on Geographical Information Science*. ScanGIS'2001, 25th-27th June 2001. Ås, Norway. pp. 3-14.